Парадигма развития науки

Методологическое обеспечение

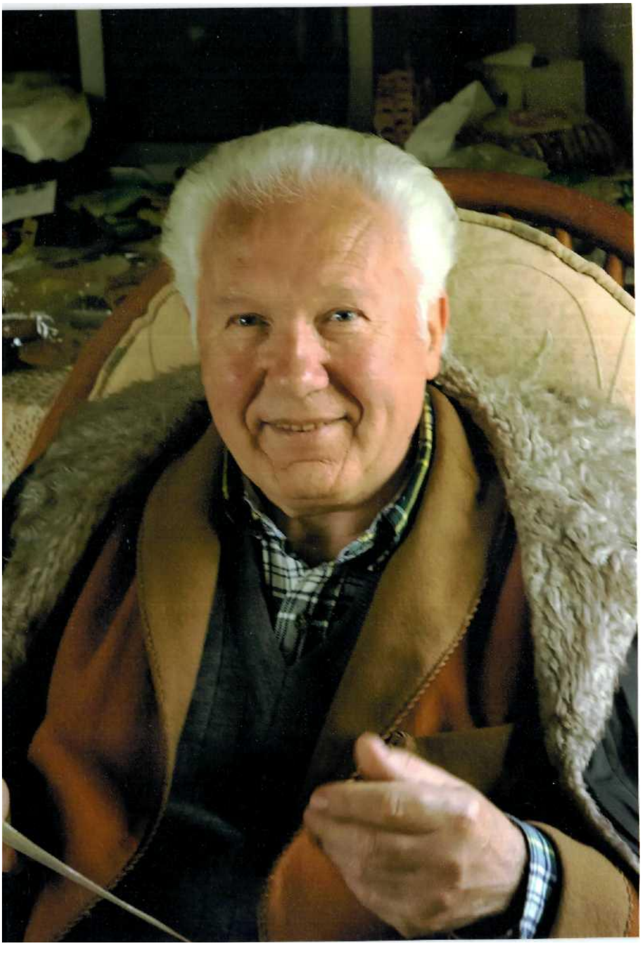
А.Е. Кононюк

ИНФОРМАЦИОЛОГИЯ

ОБЩАЯ ТЕОРИЯ ИНФОРМАЦИИ

Книга 3

Киев
«Освіта України »
2011



Кононюк Анатолий Ефимович

УДК 51 (075.8)
ББК В161.я7
К65.

Рецензент: *Н.К. Печурин* - д-р техн. наук, проф. (Нацинальный авиационный университет).

Кононюк А.Е.
К65 Информациология. Общая теория информации
К.:"Освіта України ", 2011. Книга 3, 412 с.
ISBN 978-966-7699-50-8

В настоящей работе в последовательной систематизированной форме излагаются основы информациологии как всеобъемлющей теории о естественной и искусственной информации – информации природы Вселенной и информации, созданной человеком. Дается формализация понятий информации и информационных процессов как абстрактных понятий познания и формализация самой теории информациологии. Концепция информациологии дается в аспекте социальной и природной значимости информации для отдельного человека, коллектива и общества в целом, а также – в свете информационного единства человека и природы, единства всех форм и типов информации, всех процессов информационного взаимодействия, процессов самоинформацизации Вселенной и процессов социальной (государственной) информатизации. Изложены методологические основы информациологии, базирующиеся на фундаментальном принципе информациологического подхода, на системном сочетании интеграционного и дифференцированного подходов к исследованию. Монография может быть полезной для людей самых различных специальностей – филологов, информациологов, математиков, лингвистов, правоведов, методистов, преподавателей ВУЗов, научных работников, магистров, аспирантов, докторантов, всех, кто интересуется проблемами информации, информатики, информатизации и информациологии в целом.

ББК В161.я7

ISBN 978-966-7699-50-8

Оглавление

13. Количественная оценка информации

13.1. Еще раз о понятии «информация»

Как мы уже отмечали в книге 1 «Информациология. Общая теория информации», понятие информация является одним из фундаментальных в современной науке вообще и базовым для информациологии. Информацию наряду с веществом и энергией рассматривают в качестве важнейшей сущности мира, в котором мы живем. Однако, если задаться целью формально определить понятие «информация», то сделать это будет чрезвычайно сложно.

Понятие информация в последние годы широко вошло в практику. Однако, со временем выяснилось, что в термин информация можно вкладывать различные смысловые значения. В предыдущие годы основное внимание исследователей уделялось только количественным оценкам этой важнейшей характеристики Вселенной. В то же время не менее существенно рассмотреть смысловые и ценностные аспекты этого понятия, а также учесть то, что в ряде случаев необходимо учитывать коммуникативные особенности информации. Поскольку дальнейшее продвижение в этом плане затруднено из-за неоднозначности используемых смыслов, в работе будут рассмотрены с единых позиций разные подходы к использованию термина информация, обращая внимание на все стадии информационного процесса. Рассматривается также связь понятия информации с понятиями однородности и хаоса.

Свойства мира, которые в повседневной жизни связывают с понятием *информация,* интересуют Человечество с глубокой древности. Иногда даже говорят о том, что подтверждением этого интереса служат начальные слова Евангелия от Иоанна: «В начале было слово». Тем не менее, термин информация вошел в практический обиход не очень давно. Понятия, связанные с эти термином столь важны, что он быстро получил широчайшее распространение. Быстрое и широкое распространение термина размыло границы самого понятия. В связи с этим иногда говорят о том, что информация – это чисто интуитивное, строго не определяемое понятие. Эволюция использования общих терминов всегда связана с необходимостью выделения в них разных смысловых значений. Соответственно, периодически требуется проводить их уточнение. Применительно к информации при подобном уточнении желательно ограничиться только установлением смысла разных значений термина *информация.* Вопросы же *неологизации,* то есть введения новых терминов затрагивать, видимо, преждевременно.

Если исходить из самых общих представлений, то становится очевидным, что отражение сведений об окружающем мире должно передаваться живым существам посредством некоторой функции, имеющей определенное смысловое содержание. Именно эта функция отражения свойств реального мира и рассматривается обычно в общем плане как информация или, более точно как *знание, сведения* и т.д. Нетрудно понять, что отражение мира, воспринимаемое различными сущностями, может быть разным. Поэтому такое понимание информации далеко от четких представлений. Тем не менее, это определение используется в теории познания. В теории познания также используются и другие определения понятия *информация*. Они не очень хорошо согласуются между собой. В связи с этим общий анализ понятия *информация* важен и для нучного процесса. Отметим также, что информационное отражение реального мира может воздействовать не только на живые существа, но и на некоторые машины и более простые неживые объекты. Все эти воздействия происходят в рамках того, что естественно называть *информационным процессом.*

Трудности определения понятия информация, соотносятся с двумя причинами. Первая из них очевидна. Введение нового широкого термина вычленяет широкую понятийную область. Дальнейшие работы в этой области всегда выявляют то обстоятельство, что широкие термины многоаспектны. В связи с этим после периода латентного развития возникает настоятельная необходимость в дроблении понятия и его дальнейшем уточнении. В принципе, это можно сделать как путем введения новых терминов, так и путем уточнений, достигаемых использованием различных четко установленных определений. Вторая причина обусловлена тем, что *информация – это одна из наиболее общих характеристик Универсуума (Вселенной).* Общепризнано, что Универсуум имеет иерархическую структуру. Как следствие, информация также имеет структурные уровни. Наличие структурных уровней информации требует введения уточняющих понятий и терминов. Из сказанного следует, в частности, еще одно следствие: уточнение понятия информация можно производить двумя путями. Первый путь исходит из последовательного анализа различных иерархических уровней информации. Если удачно выбрать исходный для анализа структурный уровень, то такой анализ характеризуются разумной последовательностью введения терминов. Тем не менее, трудности в выборе наиболее подходящего исходного уровня и ряд других очевидных сложностей заставляют ряд авторов в качестве отправной точки использовать идеи второго пути. Этот путь основан на общем анализе всего информационного процесса на том иерархическом уровне, который непосредственно связан с повседневной деятельностью человека. При таком подходе возможно без больших затруднений привлекать к анализу результаты из смежных областей, пользующихся понятием информации.

Понятие информации нельзя считать лишь техническим, междисциплинарным и даже наддисциплинарным термином. Информация — это фундаментальная философская категория. Дискуссии ученых о философских аспектах информации надежно показали несводимость информации ни к одной из этих категорий. Концепции и толкования, возникающие на пути догматических подходов, оказываются слишком частными, односторонними, не охватывающими всего объема этого понятия.

Попытки рассмотреть категорию информации с позиций основного вопроса философии привели к возникновению двух противостоящих концепций — так называемых, функциональной и атрибутивной. «Атрибутисты» квалифицируют информацию как свойство всех материальных объектов, т.е. как атрибут материи. «Функционалисты» связывают информацию лишь с функционированием сложных, самоорганизующихся систем.

Можно попытаться дать философское определение информации с помощью указания на связь определяемого понятия с категориями отражения и активности. *Информация есть содержание образа, формируемого в процессе отражения. Активность входит в это определение в виде представления о формировании некоего образа в процессе отражения некоторого субъект-объектного отношения.* При этом не требуется указания на связь информации с материей, поскольку как субъект, так и объект процесса отражения могут принадлежать как к материальной, так и к духовной сфере социальной жизни. Однако существенно подчеркнуть, что материалистическое решение основного вопроса философии требует признания необходимости существования *материальной среды — носителя информации в процессе такого отражения*. Итак, ***информацию следует трактовать как имманентный (неотъемлемо присущий) атрибут материи, необходимый момент ее самодвижения и***

саморазвития. Эта категория приобретает особое значение применительно к высшим формам движения материи — биологической и социальной.

Известно большое количество работ, посвященных *физической трактовке* информации. Эти работы в значительной мере построены на основе аналогии формулы Больцмана, описывающей энтропию статистической системы материальных частиц, и формулы Хартли. Соответствующие материалы по этии вопросам буду рассмотрены в настоящей работе.

Информацию следует считать особым *видом ресурса*, при этом имеется в виду толкование «ресурса» как запаса неких знаний материальных предметов или энергетических, структурных или каких-либо других характеристик предмета. В отличие от ресурсов, связанных с материальными предметами, информационные ресурсы являются неистощимыми и предполагают существенно иные методы воспроизведения и обновления, чем материальные ресурсы. В связи с таким взглядом центральными становятся следующие свойства информации: *запоминаемость, передаваемость, преобразуемость, воспроизводимость*, *стираемость*.

Подводя итог сказанному, отметим, что осуществляются (но отнюдь не завершены) усилия ученых, представляющих самые разные области знания, построить единую теорию, которая призвана формализовать понятие информации и информационного процесса, описать превращения информации в процессах самой разной природы. *Движение информации есть сущность процессов управления, которые суть проявление имманентной активности материи, ее способности к самодвижению.* С момента возникновения кибернетики управление рассматривается применительно ко всем формам движения материи, а не только к высшим (биологической и социальной). Многие проявления движения в неживых — искусственных (технических) и естественных (природных) — системах также обладают общими признаками управления, хотя их исследуют в химии, физике, механике, в энергетической, а не в информационной системе представлений. Информационные аспекты в таких системах составляют предмет новой междисциплинарной науки — ***синергетики***.

Высшей формой информации, проявляющейся в управлении в социальных системах, являются знания. Это наддисциплинарное понятие, широко используемое в общей теории познания, также претендует на роль важнейшей философской категории. *В философском плане познание следует рассматривать как один из функциональных аспектов управления*. Такой подход открывает путь к системному пониманию генезиса процессов познания, его основ и перспектив.

13.2. Энтропия как мера неопределенности выбора

Ранее отмечалось, что факт получения информации всегда связан с уменьшением разнообразия или неопределенности. В данном разделе ставятся задачи установления количественных мер неопределенности и информации и выяснения их основных свойств. Начнем рассмотрение с источника информации, который может в каждый момент времени случайным образом принять одно из конечного множества возможных состояний. Такой источник называют *дискретным источником информации.* При этом принято говорить, что различные состояния реализуются вследствие выбора их источником. Каждому состоянию источника u ставится в соответствие условное обозначение в виде знака (в частности, буквы) из алфавита данного источника: $u_1, u_2, ..., u_N$.

Для получения результата выбора источником u конкретного состояния можно высказать ряд предположений, базирующихся на априорных сведениях об источнике информации. Поскольку одни состояния выбираются источником чаще, а другие реже, то в общем случае он характеризуется *ансамблем U,* т. е. полной совокупностью состояний с вероятностями их появления, составляющими в сумме единицу:

$$U = \begin{pmatrix} u_1 & u_2 & ...u_i... & u_N \\ p(u_1) & p(u_2) & ...p(u_i)... & p(u_N) \end{pmatrix} \text{ или } U = \begin{pmatrix} u_1 u_2 ... u_i ... u_N \\ p_1 p_2 ... p_i ... p_N \end{pmatrix}, \quad (13.1)$$

причем

$$\sum_{i=1}^{N} p(u_i) = 1 \text{ или } \sum_{i=1}^{N} p_i = 1 .$$

Обе формы записи используются в дальнейшем на равных основаниях.

Опираясь на эти сведения, введем сначала меру неопределенности выбора состояния источника. Ее можно рассматривать и как меру количества информации, получаемой при полном устранении неопределенности относительно состояния источника. Мера должна удовлетворять ряду естественных условий. Одним из них является необходимость монотонного возрастания с увеличением возможностей выбора, т. е. числа возможных состояний источника N, причем недопустимые состояния (состояния с вероятностями, равными нулю) не должны учитываться, так как они не меняют неопределенности.

Ограничиваясь только этим условием, за меру неопределенности можно было бы взять число состояний, предположив, что они равновероятны. Однако такая мера противоречит некоторым интуитивным представлениям. Например, при N=1, когда неопределенность отсутствует, она давала бы значение, равное единице. Кроме того, такая мера не отвечает *требованию аддитивности,* состоящему в следующем.

Если два независимых источника с числом равновероятных состояний N и M рассматривать как один источник, одновременно реализующий пары состояний $n_i m_j$, то естественно предположить, что неопределенность объединенного источника должна равняться сумме неопределенностей исходных источников. Поскольку общее число состояний объединенного источника равно NM, то искомая функция должна удовлетворять условию

$$f(MN) = f(M)+f(N). \tag{13.2}$$

Соотношение (13.2) выполняется, если в качестве меры неопределенности источника с равновероятными состояниями и характеризующего его ансамбля U принять логарифм числа состояний:

$$H(U) = \log N. \tag{13.3}$$

Тогда при N=1 $H(U) = 0$ и требование аддитивности выполняется.

Указанная мера была предложена американским ученым Р. Хартли в 1928 г. Основание логарифма не имеет принципиального значения и определяет только масштаб или единицу неопределенности. Так как современная информационная техника базируется на элементах, имеющих два устойчивых состояния, то обычно выбирают основание логарифма равным двум. При этом единица неопределенности называется двоичной единицей или битом и представляет собой неопределенность выбора из двух равновероятных событий (*bit* — сокращение от англ. *binary digit* — двоичная единица). Если основание логарифма выбрать равным десяти, то неопределенность получим в десятичных единицах на одно состояние (дитах).

Пример 1. Определить минимальное число взвешиваний, которое необходимо произвести на равноплечих весах, чтобы среди 27 внешне неотличимых монет найти одну фальшивую, более легкую.

Общая неопределенность ансамбля U в соответствии с (13.3) составляет

$$H(U) = \log_2 27 \text{ дв. ед.}.$$

Одно взвешивание способно прояснить неопределенность ансамбля U', насчитывающего три возможных исхода (левая чаша весов легче, правая чаша весов легче, весы находятся в равновесии). Эта неопределенность

$$H(U') = \log_2 3 \text{ дв. ед.}$$

Так как

$$H(U) = 3\log_2 3 = 3H(U').$$

то для определения фальшивой монеты достаточно произвести три взвешивания.

Алгоритм определения фальшивой монеты следующий. При первом взвешивании на каждую чашку весов кладется по девять монет. Фальшивая монета будет либо среди тех девяти монет, которые оказались легче, либо среди тех, которые не взвешивались, если имело место равновесие. Аналогично, после второго взвешивания число монет, среди которых находится фальшивая, сократится до трех. Последнее, третье, взвешивание дает возможность точно указать фальшивую монету.

Предложенная мера, как мы убедились, позволяет решать определенные практические задачи. Однако она не получила широкого применения, поскольку была рассчитана на слишком грубую модель источника информации, приписывающую всем его возможным состояниям одинаковую вероятность.

Таким образом, степень неопределенности реализации состояния источника информации зависит не только от числа состояний, но и от вероятностей этих состояний. При неравновероятных состояниях свобода выбора источника ограничивается, что должно приводить к уменьшению неопределенности. Если источник информации имеет, например, два возможных состояния с вероятностями 0,99 и 0,01, то неопределенность выбора у него значительно меньше, чем у источника, имеющего два равновероятных состояния. Действительно, в первом случае результат практически предрешен (реализация состояния, вероятность которого равна 0,99), а во втором случае неопределенность максимальна, поскольку никакого обоснованного предположения о результате выбора сделать нельзя. Ясно также, что весьма малое изменение вероятностей состояний вызывает соответственно незначительное изменение неопределенности выбора.

Это позволяет сформулировать следующее требование к искомой мере неопределенности $H(p_1 \ldots p_i \ldots p_N)$: она должна быть непрерывной функцией вероятностей состояний источника $p_1 \ldots p_i \ldots p_N$ с соблюдением условия

$$\sum_{i=1}^{N} p_i = 1.$$

Наибольшее ее значение должно достигаться при равенстве вероятностей всех состояний.

Кроме того, так как мера неопределенности связывается нами только с фактом выбора, а не с множеством конкретных значений наблюдаемых явлений, то $H(p_1...p_N)$ должна быть функцией от функции распределения случайной величины и не должна зависеть от ее конкретных значений. Иначе говоря, $H(p_1...p_N)$ должна являться функционалом распределения вероятностей.

Еще одно условие состоит в том, что мера неопределенности не должна зависеть от пути выбора состояния в ансамбле. Выбор может быть как непосредственным, так и многоступенчатым. В последнем случае неопределенность выбора состояния складывается из неопределенности выбора группы состояний и неопределенностей выбора состояния в каждой группе, рассчитанных с учетом вероятности выбора данной группы:

$$H(p_1p_2 \ \dots \ p_{N-2}q_1q_2q_3q_4) = H(p_1p_2 \ \dots \ p_{N-2}p_{N-1}p_N) + \\ + p_{N-1}H(q_1/p_{N-1}, \ q_2/p_{N-1}) + p_NH(q_3/p_N, q_4/p_N), \qquad (13.4)$$

где q_1, q_2 и q_3, q_4 — вероятности состояний, образующих соответственно группы N —1 и N, причем $p_{N-1} = q_1 + q_2$ и $p_N = q_3+q_4$.

Мера неопределенности выбора дискретным источником состояния из ансамбля U, удовлетворяющая указанным условиям, была предложена американским ученым К. Шенноном. Ее называют *энтропией дискретного источника информации* или *энтропией конечного ансамбля:*

$$H(U) = -C\sum_{i=1}^{N} p_i \log p_i, \qquad (13.5)$$

где C — произвольное положительное число.

К. Шенноном высказано утверждение, а Л. Я. Хинчиным математически строго доказано, что это единственный функционал, удовлетворяющий сформулированным условиям.

Если снова ориентироваться на измерение неопределенности в двоичных единицах, то основание логарифма следует принять равным двум. Примем также C= 1. Из (13.5)

$$N(U) = -\sum_{i=1}^{N} p_i \log_2 p_i . \qquad (13.6)$$

Предложенная мера была названа энтропией не случайно. Дело в том, что формальная структура выражения (13.5) совпадает с энтропией физической системы, определенной ранее Больцманом. Согласно второму закону термодинамики энтропия H замкнутого пространства определяется выражением

$$H = -\frac{1}{M_п}\sum_{i=1}^{N} m_i \ln \frac{m_i}{M_п}, \tag{13.7}$$

где M_n — число молекул в данном пространстве; n_i — число молекул, обладающих скоростью $v_i + \Delta v$.

Так как m_i/M_n есть вероятность того, что молекула имеет скорость $v_i + \Delta v$, то (13.7) можем записать в виде

$$H = -\sum_{i=1}^{N} p_i \ln p_i.$$

Совпадение имеет глубокий физический смысл, так как в обоих случаях величина H характеризует степень разнообразия состояний системы.

Рассмотрим взаимосвязь меры К. Шеннона с мерой Хартли. Если в источнике может быть реализовано N равновероятных состояний, то вероятность каждого из них равна $p_i=(1/N)$ $(1\leq i\leq N)$ и неопределенность, по Хартли, приходящаяся на каждое состояние, выражается числом

$$H_i = \log N = -\log(1/N) = -\log p_i.$$

Будем теперь считать вероятности событий различными, а неопределенность, приходящуюся на одно конкретное состояние источника, характеризовать по аналогии величиной

$$H_i = -\log p_i. \tag{13.8}$$

Эта частная неопределенность представляет собой случайную величину, зависящую от того, какое состояние источника в действительности реализуется. Усреднив по всему ансамблю U состояний источника, найдем неопределенность, приходящуюся в *среднем на одно состояние:*

$$H(U) = -\sum_{i=1}^{N} p_i \log p_i.$$

Следовательно, мера К. Шеннона является естественным обобщением меры Хартли на случай ансамбля с неравновероятными состояниями. Она позволяет учесть статистические свойства источника информации.

Пример 2. Сравнить неопределенность, приходящуюся на букву источника информации *и* (алфавита русского языка), характеризуемого ансамблем, представленным в табл. 13.1, с неопределенностью, которая была бы у того же источника при равновероятном использовании букв.

При одинаковых вероятностях появления всех 32 букв алфавита неопределенность, приходящаяся на одну букву, составляет

$$H(U) = \log_2 32 = 5 \text{ дв. ед.}$$

Энтропию источника, характеризуемого заданным ансамблем (табл. 13.1), находим, используя формулу (13.6):

$H(U) = — 0{,}064\log_2 0{,}064 — 0{,}015\log_2 0{,}015— ... — 0{,}143\log_2 0{,}143 \approx 4{,}42$ дв. ед.

Таблица 13.1

Буква	Вероятность	Буква	Вероятность	Буква	Вероятность	Буква	Вероятность
а	0,064	й	0,010	т	0,056	ъ, ь	0,015
б	0,015	к	0,029	у	0,021	ы	0,016
в	0,039	л	0,036	ф	0,02	э	0,003
г	0,014	м	0,026	х	0,09	ю	0,007
д	0,026	н	0,056	ц	0,04	я	0,019
е, ё	0,074	о	0,096	ч	0,013	—	0,143
ж	0,008	п	0,024	ш	0,006		
з	0,015	р	0,041	щ	0,003		
и	0,064	с	0,047				

Таким образом, неравномерность распределения вероятностей использования букв снижает энтропию источника с 5 до 4.42 дв. ед.

13.3. Количественння оценка информации.

Из основной и частных задач теории информации и кодирования вытекают следующие технические и экономические проблемы: изыскание способов передачи информации при возможно меньших временных и энергетических затратах; повышение достоверности передачи, т. е. изыскание средств защиты систем передачи информации от влияния внешних физических факторов (помех, затуханий в линии связи, аппаратурных отказов и т. д.); определение рентабельности выбора способа передачи информации и метода кодирования.

В дальнейшем будет показано, что существует принципиальная возможность запроектировать устройство со сколь угодно малой вероятностью ошибки. Но практические последствия данного решения можно оценить, лишь предварительно ознакомившись с некоторыми понятиями и методами, позволяющими производить оценку эффективности систем связи. Для этого, прежде всего, необходимо уметь измерять количество и скорость передачи информации по каналам связи.

Попробуем проследить, от чего зависит количество информации. Для примера рассмотрим известную задачу о передаче места нахождения фигуры на шахматной доске. Ее можно решить одним из двух способов: закодировать каждую клетку или передать номера вертикали и горизонтали. В первом случае необходимо иметь код, отображающий 64 знака (в двоичной системе его комбинации будут состоять из шести элементов), а во втором случае — код, отображающий всего восемь знаков (в двоичной системе его комбинации будут состоять из трех элементов), но для передачи места нахождения фигуры потребуется два сообщения (номер по горизонтали и номер по вертикали). Количество информации, переданное тем и другим способом, будет одинаково. Следовательно, *количество информации, содержащейся в источнике информации, не зависит от способа ее передачи.* Длина сообщения для передачи полной информации о координате клетки также будет одинакова (6 двоичных элементов). Длина сообщений при разных способах передачи была бы различной, если бы и число качественных признаков вторичного алфавита было бы различным. Например, если число дискретных качественных признаков вторичного алфавита — 64, то длина любого сообщения о координате клетки будет состоять из одной дискретной посылки. Если для второго способа передачи информации в рассмотренном выше примере будет использовано два качественных признака, то, как мы уже знаем, длина сообщения будет равна шести, если же во вторичном алфавите будет восемь качественных признаков, то длина всего сообщения будет равна двум дискретным посылкам и т. д. Естественно, количество информации во всех случаях, когда речь будет идти о координате одной и той же клетки, будет одно и то же.

Количество информации будет увеличиваться при увеличении числа сообщений, несущих информацию о координатах новых клеток, т. е. будет удовлетворять *условию аддитивности:* при передаче сообщений одним и тем же методом, одной и той же аппаратурой, по одному и тому же каналу связи количество информации тем больше, чем большее число символов мы передаем. Действительно, чем больше слов в телеграмме, тем естественнее ожидать от нее большего количества информации; чем больше количество строк в телевизионной развертке, тем выше качество изображения; чем больше частных значений исходной функции будет передано, тем точнее она может быть воспроизведена. Не следует при этом путать термины количество и ценность информации. С точки зрения теории информации в телеграмме «Бабушка здорова. Целую, Федя» информации может быть больше или меньше, чем в телеграмме «На Марсе есть жизнь» в зависимости от того, в каком из сообщений было больше исходной неопределенности для адресата.

Способ получения сообщения влияет на количество принятой информации. Неопределенность относительно температуры утюга может быть снята как путем прикосновения пальца к утюгу, так и путем передачи на расстояние соответствующего кодированного сообщения. То же сообщение можно передать по каналу связи с помехами и без помех и т. д. При этом возможна различная полнота информации. Однако способ передачи — получения информации не может служить характеристикой последней, так как количество информации, содержащееся в ее источнике, не зависит от способов кодирования и передачи.

Количество информации I вычисляют как произведение устраненной неопределенности H, снимаемой одним сообщением, на число сообщений k. Так как сообщением является буква первичного алфавита, то k можно рассматривать как число символов первичного алфавита.

Мерой неопределенности H в теории информации является энтропия.

Если исходный ансамбль сообщений A может быть представлен конечным множеством символов абстрактного алфавита $(a_1, a_2, ..., a_i)$ с распределением вероятностей $(p_1, p_2, ..., p_i)$, то энтропия есть величина, характеризующая источник сообщений в целом и представляет собой среднюю неопределенность появления на выходе источника одного из сообщений исходного ансамбля. Но, так как исходный ансамбль может быть выражен через символы некоторого абстрактного алфавита, который мы назвали первичным, то справедливо следующее определение.

Определение 1. *Энтропия* представляет собой удельную неопределенность на символ первичного алфавита и характеризует алфавит в целом.

Определение 2. *Информацию* несут в себе те сообщения, которые снимают неопределенность, существовавшую до их поступления.

Эти сообщения могут представлять собой как результат отдельного или группы опытов, так и буквально принятые сообщения.

Информация всегда есть результат разности априорной и апостериорной энтропии.

В самом упрощенном виде это можно понимать в том смысле, что до опыта (априори) была некоторая неопределенность его исхода H_1. После опыта (апостериори) стала неопределенность H_2. Чем меньше апостериорная неопределенность, т. е. чем больше разность между H_1 и H_2, тем большее количество информации было получено.

Например, чтобы усвоить тезис, что информация есть *разность энтропии,* рассмотрим элементарный опыт, целью которого является определение пола первого вышедшего из троллейбуса человека. Троллейбус, в этом случае, является источником сообщений *A*. Исходный ансамбль сообщений в первичном алфавите может быть представлен как *A* {м, ж}, тогда априорная неопределенность этого источника сообщений *H (A)*. Если произойдет событие *B* — из двери троллейбуса выйдет человек, то апостериорная неопределенность может быть представлена как *H* (*A/B*), т. е. информацию, содержащуюся в источнике *A*, мы можем получить при свершении события *B*. Таким образом, после того как из двери троллейбуса выйдет человек, апостериорная неопределенность станет равной нулю и в результате опыта получим количество информации, равное собственной информации, содержащейся в одном из событий исходного ансамбля сообщений. Это событие можно представить в виде человека, который первым появится на выходе троллейбуса, для нашего опыта являющегося источником сообщений, но мы об этом еще не знаем, так как не произошло событие *B* (дверь еще не открылась).

Аналогичные рассуждения могут быть экстраполированы и на все подобные опыты, имеющие только два равновероятных исхода (например, бросание монеты). Во всех подобных случаях апостериорная неопределенность равна нулю. Но, если бы после бросания монеты она стала на ребро или у троллейбуса сломалась и не открылась дверь, то апостериорная неопределенность равнялась бы априорной и количество полученной информации равнялось нулю, так как не выполнилось главное условие: не произошло событие *B*. В этом случае $H(A/B) = H(A)$ и

$$I\,(A,\,B) = H(A) - H(A) = 0.$$

Если обозначить $A\{a_1,\, a_2,\, ...,\, a_m\}$, исходный ансамбль сообщений в первичном алфавите; $B\,\{b_1,\, b_2,\, ...,\, b_m\}$ множество кодовых слов, с помощью которых мы кодируем символы ансамбля *A* во вторичном алфавите, то *H(A/B)* может быть результат неоднозначного кодирования или результат потерь информации при передаче по каналу связи. В общем случае *H(A/B)* можно интерпретировать как количество информации, недостающей наблюдателю для полного снятия неопределенности, оставшейся после того, как этот наблюдатель установил, что произошло событие *B*.

Обозначение *H(A/B)* применяется для условной энтропии, которая будет рассмотрена далее. Там же подробно изложены способы вычисления *H(A/B)*. Таким образом (пока без строгих доказательств) средняя взаимная информация *I(A, B)* равна разности между энтропией *A* и условной энтропией *A* при заданном *B:*

$$I(A, B) = H(A) - H(A/B). \qquad (13.9)$$

Если ансамбль B составлен так, что осуществляется однозначное соответствие $a_1 \leftarrow b_1$; $a_2 \leftarrow b_2$; ...; $a_m \leftarrow b_m$ и при передаче не происходит потерь в канале связи, то приход любой буквы ансамбля B снимает всякую неопределенность относительно того, что именно передавалось. Тогда неопределенность вида $H(A/B)$ равна нулю. В этом случае при независимых составляющих ансамбля

$$I_B = kH(A). \qquad (13.10)$$

Казалось бы, что информация и неопределенность должны быть связаны обратной зависимостью. Но информацию мы рассматриваем как меру снятой неопределенности. С приходом каждого нового сообщения (символа первичного алфавита) общая неопределенность, существовавшая до его поступления, в среднем уменьшается на величину энтропии, присущей этому алфавиту.

Энтропия есть функция вероятности p. Чем больше вероятность события, тем меньше неопределенность относительно того, произойдет оно или нет. Нулевая или близкая к нулю вероятность может рассматриваться как большая вероятность того, что событие не произойдет. В этом случае также не будет особой неопределенности относительно факта свершения события $p = 1$ и $p = 0$.

На данном этапе изложения мы еще не можем обосновать закон, описывающий изменение энтропии в интервале от $p = 0$ до $p = 1$. С полной уверенностью можно лишь утверждать, что максимум этого значения для двух событий будет в точке $p_1 = p_2 = 0{,}5$, т. е. когда наиболее трудно предсказать результат свершения любого из них. Очевидно также, что увеличение вероятности любого из двух априорно равновероятных событий будет уменьшать общую неопределенность. При $p_1 \to 1$; $p_2 \to 0$ (и наоборот), $H \to 0$ в обоих случаях. Приведенные рассуждения позволяют предполагать, что $H=f(p)$ и функция эта выпуклая. С другой стороны, эти же рассуждения делают очевидным тот факт, что любое отклонение от равновероятного состояния, независимо от числа событий, уменьшают энтропию $H(A)$, где $A \equiv \{a_1, ..., a_m\}$ — алфавит, при помощи которого эти события отображаются. А это, в свою очередь, позволяет сформулировать следующее утверждение.

Утверждение 1. Для априорно равновероятных событий количество информации, которую мы получаем в результате свершения этих событий, уменьшается с увеличением вероятности любого из событий.

Следствием утверждения является то, что информация, заключенная в данном событии A, находится в обратно пропорциональной функциональной зависимости от вероятности $p(A)$:

$$I_A = f\left(\frac{1}{p_A}\right). \tag{13.11}$$

Выражение (13.11) справедливо при условии, что уменьшение вероятности свершения события *А* рассматривается как увеличение вероятности несовершения этого события *p'A,* тогда (13.11) может выглядеть так:

$$I_A=f(p'_A),$$

при этом обязательно $p_A + p'_A = 1$.

Теорема 1. *Если информация является функцией вероятности, то эта функция может быть только логарифмической.*

Доказательство. Согласно теории вероятностей, если $a_1, a_2, ..., a_m$ — попарно несовместимые случайные события, то вероятность появления одного из них равна сумме их вероятностей:

$$p\ (a_1, a_2, ... a_m) = p\ (a_1) + p\ (a_2) + \ldots + p(a_m),$$

но только логарифмическая функция удовлетворяет условию

$$f\ [p\ (a_1)\ p\ (a_2) \ldots p\ (a_m)] = f\ [p\ (a,)] + f\ [p\ (a_2)] + \ \ldots + f\ [p(a_m)],$$

так как логарифм произведения равен сумме логарифмов сомножителей.

При выводе логарифмической меры количества информации Р. Хартли мог пользоваться следующими рассуждениями.

Число сообщений *N,* которое можно получить, комбинируя *т* символов алфавита по *п* элементов в сообщении,

$$N = m^n. \tag{13.12}$$

Например, комбинируя два символа, можно передать восемь сообщений при $n = 3$, шестнадцать — при $n = 4$, тридцать два — при $n = 5$ и т. д. Таким образом, число сообщений *N,* а вместе с ним и количество передаваемой информации находятся в экспоненциальной зависимости от количества элементов в сообщении. Поэтому *N* нельзя непосредственно использовать как меру количества информации.

Р. Хартли предложил в качестве меры количества информации принять логарифм числа возможных последовательностей символов:

$$I== \log N = \log m^n = n \log m. \tag{13.13}$$

Основание логарифма зависит от выбранной единицы количества информации. В выражениях, где могут быть использованы произвольные логарифмы, основание логарифма не ставится.

Такое представление меры количества информации соответствует требованию аддитивности, отражает экспоненциальный характер зависимости количества возможных кодовых комбинаций от количества символов в исходном алфавите, согласуется с основным психофизиологическим законом Вебера — Фехтнера $S=K \log E$ и почти совпадает с классической формулой Больцмана для энтропии в статистической термодинамике $H_T = k \log \omega$, где S — восприятие; K — некоторая константа, зависящая от характера проводимого опыта; E — возбуждение; H_T — термодинамическая энтропия; k — константа; ω — вероятность данного состояния системы.

Подобные совпадения не случайны. Они объясняются общностью природы этих явлений; не зря во всех рассмотренных случаях наблюдается логарифмическая зависимость.

При передаче информации по каналам связи данное частное сообщение выбирают из определенного количества возможных сообщений.

Коль скоро при передаче сообщений речь идет о выборе одного из данной группы, то, очевидно, что простейшим является выбор между двумя взаимоисключающими друг друга сообщениями, априорно равновероятными для адресата. Иными словами, простейший выбор состоит в решении дилеммы «да — нет», результат решения которой может быть передан двумя качественными признаками: положительным и отрицательным импульсами, двумя частотными посылками, импульсом и паузой — в общем случае сигналами 0 и 1. Количество информации, переданное в этом простейшем случае, принято считать единицей количества информации. Так как единица количества информации представляет собой выбор из двух равновероятных событий, то это двоичная единица или *бит* (от английских слов *binary didgit* — двоичная единица).

При решении таких задач, как: будет ли четной сумма цифр взятого в трамвае билета, выйдет ли из метро первым мужчина или женщина, выпадет ли при бросании монеты орел или решка,— количество полученной информации будет равно одной двоичной единице, потому что в каждой из рассмотренных задач имеем дело с выбором из двух равновероятных событий. При решении задачи, придет ли завтра студент Иванов на лекцию или нет, информация в 1 бит может быть получена только в том случае, если до этого Иванов систематически пропускал ровно 50% лекций, так как именно в этом случае существует максимальная неопределенность опыта. Если же этот студент систематически посещал все лекции, то в сообщении, что он и завтра придет на лекцию, особой новизны не будет, количество информации будет меньше, чем 1 бит.

Если выбранная единица количества информации 1 бит есть двоичная единица, то представляется оправданным количество информации определять при помощи двоичного логарифма числа возможных последовательностей символов $I = \log_2 N$ (если бы на практике наиболее часто встречалось десять равновероятных выборов, то за единицу количества информации удобнее было бы выбрать десятичную единицу дит, а основанием логарифма в выражении (13.13) взять 10). Из свойства логарифмов известно, что $\log_b a \log_a M = \log_b M$. Это позволяет записать

$$M = 2^{\log_2 M} = 10 \log M,$$

отсюда

$$\log_2 M = \frac{\lg M}{\lg 2} = 3{,}32 \lg M.$$

Если при вычислении информации удобно пользоваться натуральными логарифмами, то единицей количества информации будет одна натуральная единица (1 нат). Натуральная единица связана с двоичной единицей следующим соотношением: 1 нат/символ = 1,443 бит/символ.

Количество информации на одну букву алфавита, состоящего из восьми равновероятных букв, $I = \log_2 8=3$ бит/символ, т. е. мера снятой неопределенности при выборе одной буквы из данного алфавита равна трем двоичным единицам. Физически это означает, что достаточно сделать три раза выбор «Да» или «Нет» (0 или 1), чтобы выбрать любой символ данного алфавита.

В этом легко убедиться, выразив символы алфавита и их аналоги в двоичном коде и произведя выбор по принципу, иллюстрированному рис. 13.1.

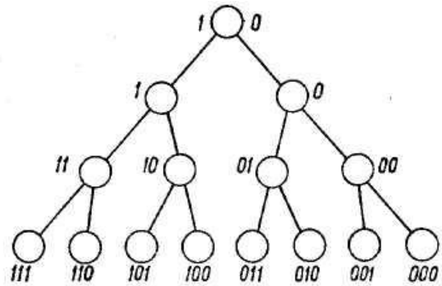


Рис. 13.1. К иллюстрации соответствия длины кода числу равновероятных выборов

Как видим, первые ветви соответствуют первому выбору между 0 и 1, причем все кодовые комбинации, у которых первый символ 1, остаются слева, а те, у которых первый символ 0,— справа. Затем, в каждой из групп производят аналогичный выбор, причем слева остаются кодовые комбинации, у которых второй символ 1, а справа —у которых второй символ 0. После третьего выбора будет однозначно определена любая буква кода, состоящего из восьми символов. Если продолжить это кодовое дерево, то можно убедиться, что передача любого символа 16-значного кода требует четыре выбора, 32-значного — пять и т. д.

Построение кода с основанием m (двоичные коды являются лишь частным случаем) эквивалентно поиску по m-значному дереву (кодовые слова представляют собой последовательность качественных признаков, которые встречаются по пути от вершины дерева к основанию). При приеме — передаче кодовых слов приход каждого символа уменьшает в m раз неопределенность того, какое из исходного множества кодовых слов было передано (имеется в виду равновероятный первичный алфавит). Общее уменьшение неопределенности находится в логарифмической зависимости от числа символов в сообщении. Информационная нагрузка на символ сообщений, передаваемых кодами равной длины, возрастает с ростом основания кода. Если принять три сообщения длиной в три символа с основанием, соответственно, 2, 3 и 10, то:

а) для первого сообщения приход первого символа исключает четыре варианта из восьми, второго символа — два из четырех, третьего — один из двух возможных;

б) для второго сообщения приход первого символа исключает 18 вариантов из 27 возможных, второго — шесть из девяти, третьего — два из трех;

в) для третьего сообщения первый символ исключает 666 вариантов из 1000; второй — 224 из 334; третий — девять из десяти.

Первое сообщение тремя символами сняло неопределенность о возможности получения одного из восьми сообщений; второе — одного из 27; третье — при той же длине сообщения сняло неопределенность о возможности получения одного из 1000 сообщений. Очевидно, что информационная нагрузка на символ этих сообщений — разная.

Приведенные выше соображения позволяют сформулировать следующие утверждения.

Утверждение 2. Не существует двух различных единиц измерения информации, переданной одним и тем же кодом.

Утверждение 3. Единица измерения количества принимаемой информации равна основанию кода, при помощи которого отображаются принимаемые сообщения.

Из последнего утверждения следует, что единиц измерения информации может быть столько, сколько существует оснований кодов, передающих ее.

Основание логарифма при измерении энтропии должно быть равно единице измерения количества информации.

Логарифм с натуральным основанием следует применять для вычисления энтропии непрерывных сообщений, больцмановской энтропии замкнутой термодинамической системы в общем случае там, где необходимо подчеркнуть экспоненциальный характер энтропии, где исходное разнообразие не может быть определено целым числом.

Если число передаваемых символов первичного алфавита m_1 является одним из чисел степенного ряда m_2^n, то количество информации определяется степенью m_2.

В многопозиционных кодах (с основанием больше 2) при $m \geq 3$ применение единицы количества информации бит недопустимо без соответствующего модуля перехода от системы логарифмов с основанием m в выбранную систему логарифмов, например, с основанием 2. В этом случае бит может быть универсальной единицей измерения количества информации наиболее удобной и с точки зрения ее восприятия, и с точки зрения ее соответствия характеру современной техники.

Вернемся теперь к понятию энтропии, которая представляет собой среднюю неопределенность на опыт, событие, сообщение, состояние и т. д. При этом будем иметь в виду более общий случай, когда первичный алфавит представляет элементы систем, события, языковые алфавиты с неравновероятным законом распределения.

Говорить об энтропии i-го опыта, состояния, букв для неравновероятных алфавитов нельзя, если буква, состояние, опыт представляет собой одну из букв, состояний, опытов, составляющих в вероятностном отношении полную группу, так как энтропия есть величина, которая характеризует весь алфавит. Сама по себе величина $\log p_i$ не говорит ни о чем, кроме того, что мы имеем дело с некоторым неравновероятным событием, сообщением, в общем случае с буквами некоторого неравновероятного абстрактного алфавита $A\{a_1, a_2, \ldots, a_m\}$. В вероятностном отношении события, представляемые алфавитом A, должны составлять полную группу, т. е. если они попарно несовместимы, то при каждом повторении испытания должно произойти хотя бы одно из них. С другой стороны, если события $a_1, a_2, \ldots a_m$ образуют полную группу, то $p(a_1) + p(a_2) + \ldots + p(a_m) = 1$. Поэтому нельзя рассматривать отдельное событие с вероятностью $0 < p(a_i) < 1$. Энтропии отдельно взятого 1-го события с вероятностью $0 < p(a_i) < 1$ не существует. Говоря об энтропии неравновероятного алфавита, нельзя не учитывать вероятностную взаимосвязь каждой буквы алфавита со всеми остальными буквами, составляющими полную группу событий. Поэтому энтропия на букву любого алфавита должна рассматриваться как математическое ожидание **М** величины — $\log p(a_i)$:

$$H(A) = \mathbf{M}[-\log p(a_i)] = -\sum_{i=1}^{m} p(a_i) \log p(a_i). \tag{13.14}$$

Докажем справедливость выражения (13.14).

Теорема 2. *Энтропия дискретного эргодического источника сообщений равна математическому ожиданию функции распределения ξ вероятностей появления на выходе источника сообщений исходного ансамбля.*

Доказательство. Если алфавит источника сообщений содержит k_1 символов a_1, k_2 символов a_2, ..., k_i символов a_i, ... k_m символов a_m, то общее число элементов исходного множества

$$k = k_1 + k_2 + \cdots + k_i + \cdots + k_m = \sum_{i=1}^{m} k_i. \tag{13.15}$$

Вероятность появления в сообщении символов с i-м признаком равна p_i, вероятность совместного появления символов исходного множества, имеющих вероятность

для $a_1 - p_1^{k_1}$, для $a_2 - p_2^{k_2}$, ..., для $a_m - p_m^{k_m}$ равна

$$p = \prod_{i=1}^{m} p_i^{k_i}. \tag{13.16}$$

При бесконечном числе передач N сообщений в алфавите A величина k_i/k будет характеризовать вероятность появления символа a с i-м признаком, т. е. при $N \to \infty$ $k_i/k \to p_i$, что позволяет не привязываясь в обозначениях к конкретному алфавиту, записать

$$k_i \approx k_{p_i}. \tag{13.17}$$

Подставим значение k_i в (113.16), тогда

$$p = \prod_{i=1}^{m} p_i^{k_{p_i}}. \tag{13.18}$$

Учитывая (13.11) и (13.18), выражение для I может быть представлено следующим образом:

$$I = f\left(\frac{1}{\prod_{i=1}^{m} p_i^{k_{p_i}}}\right). \tag{13.19}$$

Поскольку $f\ (p)$ согласно теореме 1 есть функция логарифмическая, то

$$I = \log\left(\frac{1}{\prod_{i=1}^{m} p_i^{kp_i}}\right) = -\log \prod_{i=1}^{m} p_i^{kp_i} =$$

$$= -kp_1 \log p_1 - kp_2 \log p_2 - \cdots - kp_m \log p_m = -k \sum_{i=1}^{m} p_i \log p_i.$$

Для случая однозначного кодирования, отсутствия взаимозависимости между символами и отсутствия помех в канале связи $I = kH$, а

$\sum_{i=1}^{m} p_i \log p_i -$ есть математическое ожидание величины $\log p_i$. Следовательно,

$$H = M[-\log p_i] = -\sum_{i=1}^{m} p_i \log p_i. \quad (13.20)$$

Соотношение (13.20) было получено Шенноном для определения среднего количества информации в сообщении с произвольными вероятностями появления значений символов. При равновероятных символах, т. е. при $p_i = 1/m$, формула Шеннона переходит в формулу Хартли:

$$I = -k \sum_{i=1}^{m} p_i \log p_i = -k \sum_{i=1}^{m} \frac{1}{m} \log \frac{1}{m} = k \log m. \quad (13.21)$$

Выражение (13.21) несколько отличается от (13.14). Для разъяснения этого обстоятельства вернемся к понятиям первичного и вторичного алфавита и с помощью примера рассмотрим взаимосвязь между числом качественных признаков первичного и вторичного алфавитов.

Пример 1. Тексты, составленные из 32 букв украинского алфавита, передаются по телетайпу при помощи двух качественных признаков: наличия и отсутствия токовой посылки. Чему равно количество информации, приходящееся на одну принятую букву, на k принятых букв?

Решение. Число качественных признаков первичного алфавита $m_1 = 32$; число качественных признаков вторичного алфавита $m_2 = 2$. Для передачи 32 букв при помощи двух качественных признаков их следует комбинировать по пять символов в сообщении, так как

$$m_1 = m_2^n = 2^5 == 32,$$

отсюда длина сообщения во вторичном алфавите $n = 5$.

Количество информации на букву относительно первичного алфавита

$$I = \log_2 m_1 = \log_2 32 = 5 \text{ бит}.$$

Количество информации на букву относительно вторичного алфавита

$$I = \log_2 m_2^n = \log_2 2^5 = 5 \log_2 2 = 5 \text{ бит.}$$

Количество информации на k принятых букв

$$I = k \log_2 m_1 = kn \log_2 m_2. \qquad (13.22)$$

Выражение (13.22) подтверждает отсутствие противоречий между выражениями (13.13) и (13.11).

Понятие информации связано со снятием неопределенности, которая существовала до получения сообщения. И чем большая неопределенность была до передачи сообщения, тем большее количество информации содержится в принятом сообщении.

Предположим, в урне находится 10 000 шаров: 9999 черных и один белый. Вынимаем шар, передаем сообщение о цвете шара в какой-то условный пункт приема, кладем шар обратно в урну, перемешиваем щуры и повторяем процедуру. Вероятности получения сообщения о том, что вынут черный или белый шар, будут равны соответственно $p_1 =$ 0,9999 и $p_2 =$ 0,0001. Количество передаваемой информации при этом

$$I = -(p_1 \log_2 p_1 + p_2 \log_2 p_2) = -(0{,}9999 \log_2 0{,}9999 + 0{,}0001 \log_2 0{,}0001) =$$
$$= 0{,}001329 + 0{,}000144 = 0{,}001473 \text{ бит/опыт.}$$

Следовательно, заранее предсказать, какого цвета будет вынут шар, не представляет особого труда, так как мы почти уверены в результате опыта. Количество получаемой информации ничтожно мало. Если мы заранее уверены в ответе, то ответ несет нулевую информацию.

Если в примере с шарами в урне было бы 5000 черных и 5000 белых, то трудно было бы предсказать содержание сообщения. В этом случае неопределенность максимальна, так как вероятности появления черного и белого шаров равны друг другу: $p_1 = p_2 = 0{,}5$. Количество полученной информации при этом

$$I = -(p_1 \log_2 p_1 + p_2 \log_2 p_2) = -(0{,}5 \log_2 0{,}5 + 0{,}5 \log_2 0{,}5) =$$
$$= 1 \text{ бит/опыт.}$$

Чем больше априорная неопределенность, тем большее количество информации получается при снятии ее. В этом смысле неопределенность является удобной мерой оценки количества информации при исследовании ее свойств.

После опыта, когда решен вопрос о том, какой символ (или группа символов) передан, в каком состоянии находился элемент системы (или вся система), неопределенность становится равной нулю. Таким образом, если при передаче информации не было информационных потерь, то в случае однозначного кодирования количество информации на символ сообщения будет точно равно H, а количество информации при передаче k символов $I = kH$.

Следует различать понятия «количество информации» и «объем информации».

Количество информации уменьшается с числом повторений по экспоненте как логарифм от числа повторений.

Объема информации, как такового, не существует. Речь может идти о количестве принятых знаков, составляющих числовые данные, тексты... Именно в этом смысле часто употребляют термин «объем информации».

Следовательно, «объем информации» не зависит от числа повторений (при повторном чтении одного и того же текста его информативность быстро падает, а число знаков остается прежним).

Количество информации вычисляется относительно первичного алфавита, а «объем информации» — относительно вторичного алфавита.

Количество информации зависит от вероятностных характеристик первичного алфавита, а объем не зависит. «Объем информации» зависит от длины сообщения во вторичном алфавите *n и* равен

$$Q = kn.$$

Пример 2. Определить объем и количество информации при передаче русского текста из 350 букв при помощи пятизначного двоичного кода.

Решение: Известно, что энтропия русского алфавита без учета взаимозависимости между буквами, равна 4,358 бит/букву. Тогда количество информации

$$I = kH = 350 \bullet 4{,}358 \text{ бит/букву} = 1525{,}3 \text{ бит}.$$

Длина сообщения во вторичном алфавите $n = 5$ дв. знаков; «объем информации» $Q = kn = 350 \bullet 5$ дв. знаков = 1750 дв. знаков.

Количество информации равно объему, в соизмеримых единицах, если выполняются условия:

1) символы первичного алфавита встречаются в сообщениях с равной вероятностью; количество символов первичного алфавита является целой степенью двух, в случае, если вторичный алфавит — двоичный, и целой степенью m_2, в случае, если $m_2 > 2$;

2) для неравновероятных алфавитов $p_i = m_2^{-n_i}$, где n — целое. Во всех остальных случаях «объем информации» в соизмеримых единицах будет больше, чем количество информации.

Если $m_2 > 2$, то, при сравнении количества и «объема информации», в формуле для вычисления количества информации основание логарифма должно быть равно m_2. При этом равенство объема и количество информации при передаче равновероятных сообщений возможны лишь при условии $m_1 = m^n{}_2$. При $m_1 < m^n{}_2$ всегда $I < Q$ при $m_1 > m^n{}_2$ невозможно однозначное декодирование сообщений.

Выводы: 1. *Информация имеет количественную оценку.*

2. Зависимость между количеством информации и количеством комбинаций, составленных из данного алфавита, — логарифмическая.

3. Общепринятая единица количества информации — двоичная. Получается она при выборе одного из двух равновероятных символов и равна 1 бит.

*4. **Количество информации** определяется не количеством ситуаций (хотя, безусловно, зависит от него), а устранением неизвестности о передаваемом факте.*

5. Количество информации определяется относительно первичного алфавита и не зависит от способа ее передачи и длины сообщения во вторичном алфавите.

*6. **Объем информации** определяется относительно вторичного алфавита и равен произведению средней длины сообщения во вторичном алфавите на число сообщений.*

7. Количество информации не может быть больше объема той же информации в соизмеримых единицах.

*8. **Информация** есть единичные сообщения либо их совокупности, которые снимают неопределенность, существующую до их поступления.*

9. Энтропию появления отдельно взятой буквы алфавита определить нельзя, так как энтропия характеризует алфавит в целом и представляет собой математическое ожидание величины — $\log p_i$. *Поэтому оперировать понятием «энтропия» относительно появления кодовых слов на выходе источника сообщений можно лишь в том смысле, что неопределенность появления кодового слова на выходе источника сообщений равна энтропии источника.*

13.4. Безусловная энтропия и ее свойства

Безусловная энтропия — удельное количество информации на один элемент сообщений, составленных из алфавитов, между символами которых не наблюдается взаимозависимость. Когда говорят о безусловной энтропии, определение «безусловная» опускают. Оговариваются другие виды энтропии, как менее распространенные.

Рассмотрим понятие энтропии как меры неопределенности некоторого опыта, исход которого зависит от выбора одного элемента из множества исходных. Множество исходных элементов называется *выборочным пространством.* Вероятности нахождения элементов исходного множества в том или ином состоянии есть числа положительные, а сумма их равна единице (в противном случае результат опыта не относился бы к полной группе событий).

Выборочное пространство и его вероятностные характеристики представляют собой *ансамбль сообщений.* Для дискретного ансамбля вероятность события равна сумме вероятностей элементов выборочного пространства, содержащихся в этом событии.

Ансамбль сообщений на выходе источника будем называть *ансамблем источника сообщений* и обозначать буквой *A.* Абстрактный алфавит, при помощи которого мы представляем исходное множество элементов источника сообщений, обозначается $\{a_1, a_2, ..., a_i, ..., a_m\}$. Вероятности появления буквы на выходе источника сообщений обозначим

$$p(a_1), p(a_2), \ldots, p(a_i), \ldots, p(a_m), \sum_{i=1}^{m} p_{a_i} = 1.$$

В этом случае *энтропия источника сообщений*

$$H(A) = -\sum_{i=1}^{m} p(a_i) \log p(a_i)$$

представляет собой неопределенность появления на выходе источника сообщений буквы первичного алфавита.

Ансамбль сообщений на входе приемника будем называть *ансамблем приемника сообщений* и обозначать буквой *B.* Для того чтобы отличить переданные и принятые сигналы, абстрактный алфавит, в котором представлен ансамбль приемника сообщений, обозначается $\{b_1, b_2, \ldots$ $\ldots, b_j, \ldots, b_m\}$, а соответствующие вероятности — $p(b_1), p(b_2), \ldots$ $\ldots.\ p(b_j), \ldots,\ p(b_m)$.

Энтропия приемника сообщений

$$H(B) = -\sum_{i=1}^{m} p(b_j) \log p(b_i)$$

представляет собой неопределенность появления на входе приемника буквы после ее появления на выходе источника сообщений. Если в канале связи не происходит потерь информации, то всегда буква a_1 соответствует букве b_1, a_2 — b_2 и т. д. При этом $H(A) = H(B)$.

Понятие энтропии используется не только при передаче сообщений. Энтропия широко применяется для описания состояния механических и термодинамических систем, для изучения свойств алфавитов различных языков, при исследовании экономических систем и т. д.

Пример 3. Чему равна энтропия системы, состоящей:

а) из двух элементов, каждый из которых может с равной вероятностью находиться в двух состояниях?

б) из трех элементов, каждый из которых может с равной вероятностью находиться в четырех состояниях?

Р е ш е н и е: а) число элементов системы $m = 2$; число состояний

n = 2; энтропия $H = \log_2 m^n = n \log_2 m = 2 \log_2 2 = 2$ бит/состояние;

б) число элементов системы m = 3; число состояний n = 4; энтропия $H = \log_2 m = 4 \log_2 3 = 6{,}32$ бит/состояние или условный источник сообщений принимает одно из $m_1 = m_2^n = 3^4 = 81$ возможных равновероятных состояний. Энтропия на состояние

$$H = -\log_2 p = -\log_2 \frac{1}{m_1} = \log_2 m_1 = n \log_2 m_2 = 4 \log_2 3 = 6{,}32 \text{ бит/состояние.}$$

В данном примере мы рассматривали равновероятные состояния элементов системы. Поэтому для вычисления энтропии использовалась формула Хартли (13.13).

Если между элементами системы не наблюдается никаких корреляционных связей, а состояния системы (или буквы первичного алфавита, представляющего их) неравновероятны, то энтропия такой системы вычисляется по формуле Шеннона (13.20) при значениях k = 1.

Пример 4. Определить энтропию системы, состояние которой описывается дискретной величиной со следующим распределением вероятностей состояний

a_1	a_1	a_2	a_3	a_4
p_{a_i}	0,1	0,2	0,3	0,4

Решение.

$$H = -\sum_{i=1}^{m} p_i \log_2 p_i = -(0{,}1 \log_2 0{,}1 + 0{,}2 \log_2 0{,}2 + 0{,}3 \log_2 0{,}3 + 0{,}4 \log_2 0{,}4) = 1{,}846440 \text{ бит/состояние.}$$

В примерах 3 и 4 состояния элементов условных источников сообщений не зависят друг от друга. Энтропия источников, у которых элементы ансамблей сообщений взаимозависимы, называется *энтропией нулевого порядка.*

При исследовании свойств энтропии наибольший интерес представляет ее зависимость от числа m возможных признаков (качеств) и вероятности p_i появления в сообщении элемента с i-м признаком.

Свойство 1. Если известно, что событие наверняка произойдет (эквивалентно передаче сообщения с одним качественным признаком), то его энтропия равна нулю.

Доказательство. Если $m_1 = i = 1$, то вероятность приема сигнала с i-м признаком p_i = 1, тогда

$$H_i = \sum_{i=1}^{m} p_i \log \frac{1}{p_i} = 1 \log 1 = 0.$$

Это свойство хорошо воспринимается интуитивно, так как если заранее известно, что будет передано сообщение с i-м признаком, то очевидно, что при его получении ничего нового мы не узнаем, т. е. получим нулевую информацию.

Свойство 2. Если известно, что событие наверняка не произойдет (эквивалентно отсутствию *i*-го качественного признака), то его энтропия равна нулю.

Доказательство. Если вероятность появления *i*-го признака в ансамбле сообщений равна нулю, то слагаемое с этим признаком принимает вид неопределенности типа нуль, умноженный на бесконечность. Действительно,

$$p_i \log \frac{1}{p_i} = 0 \cdot \infty.$$

Раскроем эту неопределенность, используя правило Лопиталя. Для этого прежде всего неопределенность вида 0 •∞ приводим к виду ∞/∞:

$$\lim_{p_i \to 0} \left(p_i \log \frac{1}{p_i} \right) = \lim_{p_i \to 0} \left(\frac{\log \frac{1}{p_i}}{\frac{1}{p_i}} \right).$$

Обозначим

$$\frac{1}{p_i} = k \text{ (при } p_1 \to 0 \;\; k \to \infty).$$

Тогда можно записать

$$\lim_{p_i \to 0} \left(\frac{\log \frac{1}{p_i}}{\frac{1}{p_i}} \right) = \lim_{k \to \infty} \frac{\log k}{k}.$$

Согласно правилу Лопиталя,

$$\lim_{k \to \infty} \frac{\log k}{k} = \frac{d}{d_k} \left(\frac{\log k}{k} \right),$$

но производная

$$(\log k)' = \frac{1}{k} \log \mathrm{e},$$

т.е.

$$\lim_{k \to \infty} \frac{\log k}{k} = \frac{\frac{1}{k} \log \mathrm{e}}{1} = 0.$$

Известно, что если отдельные слагаемые стремятся к нулю, то к нулю стремится и сумма, т. е. окончательно можно записать

$$\lim_{p_i \to 0} \sum_{i=1}^{m} p_i \log \frac{1}{p_i} = 0.$$

Таким образом найдены два крайних условия, при которых энтропия минимальна и равна нулю.

Свойство 3. Энтропия — величина ограниченная. Для слагаемых — $p_i \log p_i$ в диапазоне $0< p_i \leq 1$ ограниченность очевидна. Остается определить предел, к которому стремится слагаемое — $p_i \log p_i$ при $p_i \to 0$, поскольку —$\log p_i$ при этом неограниченно возрастает:

$$\lim_{p_i \to 0}(-p_i \log p_i) = \lim_{p_i \to 0} \frac{\log(1/p_i)}{1/p_i}.$$

Обозначив $\alpha=1/p_i$ и воспользовавшись правилом Лопиталя, получим

$$\lim_{p_i \to 0}(-p_i \log p_i) = \lim_{\alpha \to \infty} \frac{\log \alpha}{\alpha} = \lim_{\alpha \to \infty} \frac{(1/\alpha)\log e}{1} = 0.$$

Свойство 4. Если опыт представляет из себя цепь равновероятных событий (эквивалентно передаче сообщений из равновероятных взаимонезависимых символов), то его энтропия максимальна.

Доказательство. Максимальное значение энтропии, которая представляет собой сумму вида $-\sum_{i=1}^{m} p_i \log p_i$, будет в том случае, когда значение слагаемых суммы будет максимальным. Максимальная величина слагаемого может быть найдена после его исследования на экстремум. Для этого составим вспомогательную функцию F вида

$$F = -\sum_{i=1}^{m} p_i \log p_i - \lambda, \tag{13.23}$$

где λ — множитель Лагранжа.

Используя метод Лагранжа для отыскания условного экстремума, представим (13.23) в виде

$$F = -\sum_{i=1}^{m} p_i \log p_i - \lambda \sum_{i=1}^{m} p_i.$$

Это выражение фактически не отличается от (13.23), так как

$$\sum_{i=1}^{m} p_i = 1.$$

Теперь представим F в следующем виде:

$$F = -\sum_{i=1}^{m} [p_i \log p_i + \lambda p_i] = \sum_{i=1}^{m} F_i$$

и найдем максимум этой функции, который будет при

$$\frac{dF_i}{dp_i} = -\log p_i - \log e - \lambda = 0, \qquad (i = 1,\ 2,\ \ldots,\ m).$$

$$\log p_i = -\log e - \lambda,$$

где i=1, 2, …, m.

Таким образом, максимальная величина слагаемого $p_i \log p_i$ не зависит от p_i, что может быть только при условии равенства вероятностей появления в сообщении любого качественного признака из алфавита с m качественными признаками, т. е. при

$$p_1 = p_2 = \cdots = p_m = \frac{1}{m}.$$

Сиойство 5. Энтропия с конечным числом исходов всегда положительна.

Доказательство. Выражение для энтропии представляет собой математическое ожидание правильной дроби со знаком «минус». Так как логарифмы правильных дробей всегда отрицательны, то значение выражения

$$H = \mathrm{M}[-\log p_i]$$

всегда положительное.

Итак, энтропия есть величина вещественная, положительная и имеет экстремум.

График функции — $p_i \log p_i = f(p_i)$ изображен на рис. 13.2.

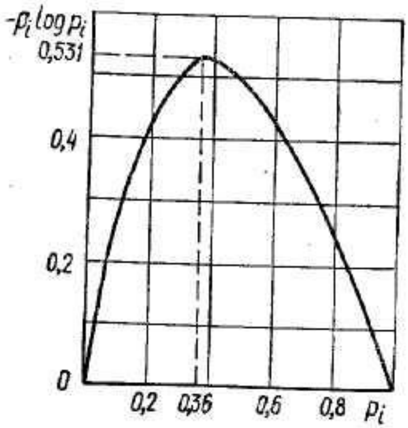


Рис. 13.2. График функции
$p_i \log p_i = f(p_i)$ при $m > 2$

Этот график представляет интерес с той точки зрения, что позволяет оценить влияние вероятности появления отдельного символа на величину выражения энтропии для сообщения в целом. Как видно из графика, при $p_i < 0{,}1$ величина $p_i \log p_i$ растет круто. Это означает, что на данном участке даже незначительное уменьшение вероятности p_i ведет к резкому уменьшению слагаемого $p_i \log p_i$, т. е. при *малых* *значениях вероятности* p_i *члены в выражении энтропии, содержащие* p_i, *не играют существенной роли и часто могут быть опущены.*

Из рис. 13.2 также видно, что наибольшие значения слагаемых вида $p_i \log p_i$ принимаются при вероятностях появления импульса с i-м признаком, лежащих в области от 0,2 до 0,6. Это понятно, так как при малых вероятностях появления i-го признака легко предсказать его отсутствие в сообщении и, наоборот, при больших вероятностях появления i-го признака легко предсказать его присутствие в сообщении. В обоих случаях величина неопределенности, существующей до получения сообщения, будет мала. Соответственно мало и количество информации при снятии этой неопределенности, что и иллюстрируется рис. 13.2.

Свойство 6. Энтропия бинарного сообщения обладает свойством симметрии: увеличение или уменьшение вероятности любого из элементов бинарного сообщения на одну и ту же величину ведет к одинаковому изменению энтропии.

Доказательство. Если число символов в сообщении равно двум,то

$$H = -\sum_{i=1}^{2} p_i \log_2 p_i = -(p_1 \log_2 p_1 + p_2 \log_2 p_2). \tag{13.24}$$

График функции, соответствующей (13.24), представляющий собой выпуклую функцию с одним экстремумом в точке $p_i = 1/m$, интересен также и тем, что полностью объясняет факт уменьшения энтропии как при увеличении, так и при уменьшении вероятности одного из событий $a_1, a_2, ..., a_i, ..., a_m$ (образующих в вероятностном отношении полную группу) относительно некоторого исходного равновероятного состояния. Если же не брать за исходное равновероятное состояние, то вообще нельзя утверждать, что увеличение вероятности ведет к уменьшению энтропии, так как если увеличение вероятности свершения одного из событий ведет к выравниванию вероятностей в группе, то энтропия будет увеличиваться.

Поскольку в (13.24)

$$\sum_{i=1}^{m} p_i = 1, \text{ то } p_1 + p_2 = 1.$$

Обозначим для удобства $p_1 = p_2 = p$, тогда $p_2 = 1 - p_1 = 1 - p$. Подставим значения p_1 и p_2 в (13.24):

$$H = -[p \log_2 p + (1-p) \log_2 (1-p)]. \tag{13.25}$$

График функции (13.25) представлен на рис. 13.3.

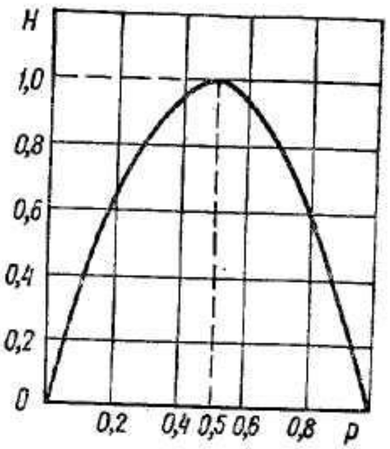


Рис. 13.3. График функции $H = f(p)$ при $m = 2$

Как видим, отклонение от величины $p = 0{,}5$ в любую сторону ведет к одинаковому уменьшению энтропии.

Итак, энтропия бинарного сообщения изменяется от 0 до 1 и достигает максимума при равных вероятностях появления в сообщении признаков, т. е. при $p_1 = p_2 = 0{,}5$.

Свойство 7. Энтропия источника u с двумя состояниями u_1 и u_2 изменяется от нуля до единицы, достигая максимума при равенстве их вероятностей:

$$p(u_1) = p = p(u_2) = 1 — p = 0{,}5.$$

График зависимости $H(U)$ в функции p

$$H(U) =-[p \log p + (1-p) \log (1-p)]$$

приведен на рис. 13.4.

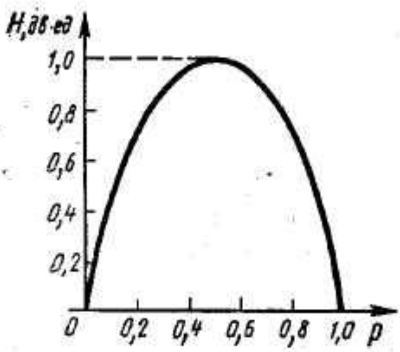


Рис. 13.4

При p□ (1—p) частная неопределенность, приходящаяся на состояние u_1, велика, однако такие состояния источника весьма редки. Состояния u_2 реализуются часто, но неопределенность, приходящаяся на такое состояние, очень мала. Поэтому энтропия, характеризующая среднюю неопределенность на одно состояние ансамбля, также мала. Аналогичная ситуация наблюдается при p□ (1—р).

Отметим, что энтропия непрерывно зависит от вероятностей отдельных состояний, что непосредственно вытекает из непрерывности функции — $p \log p$.

Свойство 8. Энтропия сложного сообщения, состоящего из некоторых взаимонезависимых частных сообщений, равна сумме энтропии, составляющих его сообщений.

Доказательство. Предположим, имеются ансамбли сообщений $A\ \{a_i\}$ и $B\ \{b_j\}$ с энтропиями, соответственно $H\ (A)$ и $H\ (B)$. Для независимых событий вероятность совместного события AB равна произведению вероятностей событий A и B, то, обозначив через p_i, p_j и $p_{i,j}$ вероятности событий соответственно A, B и AB и используя формулу для энтропии случайного события (13.21), выражение для энтропии сообщения AB запишем следующим образом:

$$H(AB) = -\sum_{i,j} p_{i,j} \log p_{i,j} = -\sum_{i,j} p_i p_j \log p_i p_j =$$

$$= -\sum_{i,j} p_i p_j (\log p_i + \log p_j) = -\sum_i p_i \log p_i \sum_j p_j - \sum_j p_j \log p_j \sum_i p_i,$$

так как

$$\sum_i p_i = 1 \text{ и } \sum_j p_j = 1.$$

Окончательно можно записать

$$H(AB) = -\sum_i p_i \log p_i - \sum_j p_j \log p_j = H(A) + H(B),$$

$$H(AB) = -\sum_{i,j} p_{i,j} \log p_{i,j} = -\sum_{i,j} p_i p_j \log p_i p_j =$$

$$= -\sum_{i,j} p_i p_j (\log p_i + \log p_j) = -\sum_i p_i \log p_i \sum_j p_j - \sum_j p_j \log p_j \sum_i p_i,$$

что и требовалось доказать.

Иллюстрацией данного свойства может служить следующий пример.

Послан запрос относительно даты определенного мероприятия, которое должно состояться в будущем году. В ответ сообщили, что оно состоится 5 сентября. До получения этого сообщения адресат, посылавший запрос, мог с равной вероятностью предположить любой из 365 дней будущего года. Это значит, что неопределенность его первоначальных представлений о дате предстоящего мероприятия составляла в соответствии с формулой энтропии:

$$H_{\Sigma} = - \sum_{i=1}^{i=365} \frac{1}{365} \log_2 \frac{1}{365} = 8{,}54 \text{ бит.}$$

Полученное сообщение («5 сентября») полностью устранило эту неопределенность, следовательно, количество содержащейся в нем информации составляло

$$I_{«5IX»} = 8{,}54 \text{ бит.}$$

Используя формулу энтропии, нетрудно определить, что в сообщенном слове «сентябрь» в данном случае содержится количество информации, равное

$$I_{«IX»} = 3{,}6 \text{ бит.}$$

Принцип такого расчета вполне очевиден: полученное сообщение позволило получателю отметить один из 12 месяцев. До получения сообщения вероятности всех 12 месяцев были равны и составляли

$$p_1 = p_2 = \ldots p_{12} = \frac{1}{12}.$$

При этом первоначальная неопределенность относительно месяца проведения интересующего мероприятия составляла:

$$H_2 = - \sum_{i=1}^{i=12} \frac{1}{12} \log_2 \frac{1}{12} = 3{,}6 \text{ бит.}$$

Ответное сообщение полностью устранило данную неопределенность, поэтому в рассматриваемом конкретном примере

$$I_{«IX»} = |H_2| = 3{,}6 \text{ бит.}$$

Аналогично определяется количество информации, содержащейся в сообщенном числе «5», позволившем получателю отметить один из 30 дней сентября:

$$I_{«5»} = - \sum_{i=1}^{i=30} \frac{1}{30} \log_2 \frac{1}{30} = 4{,}93 \text{ бит.}$$

Сопоставление полученных значений $I_{«5IX»}$, $I_{«IX»}$, $I_{«5»}$ может служить иллюстрацией свойств аддитивности информации: общее количество информации, содержащейся в названной дате «5 сентября», равно сумме информации, заключенной в названии месяца и в сообщенном числе: $I_{«5IX»} = I_{«IX»} + I_{«5»}$. (Несовпадение значений $I_{«5IX»}$ и $I_{«IX»} + I$ в третьем знаке обусловлено тем, что не все месяцы имеют ровно 30 дней и поэтому $I_{«5IX»}$ не точно равно $\frac{1}{30} \cdot \frac{1}{12} = \frac{1}{360}$ $\left(p_{«5\ IX»} = \frac{1}{365}\right)$.

Как уже отмечалось, аддитивность вычисленных значений информации $I_{«5IX»}$, $I_{«IX»}$, $I_{«5»}$ объясняется тем, что формула энтропии включает в себя логарифмическую функцию. При определении вероятности совпадепия двух независимых событий (в данном случае совпадения сообщений «сентябрь» и «5») вероятности перемножаются:

$$p_{«5IX»} = p_{«IX»} \bullet p_{«5»}$$

Благодаря тому, что логарифм произведения равен сумме логарифмов и соблюдается аддитивность:

$$\log p_{«5IX»} = \log p_{«5»} + \log p_{«IX»}.$$

В объективной реальности аддитивность является основой процессов накопления приходящей из внешней среды информации и ее сохранения в структуре различных материальных систем.

В приведенном примере полученное сообщение («5 сентября») полностью устранило неопределенность исходных сведений относительно интересующей получателя сообщения даты. Однако теории информации и технике связи чаще приходится иметь дело с сообщениями, которые устраняют неопределенность исходных сведений не полностью, а частично. В этих случаях количество содержащейся в сообщении информации определяется уменьшением неопределенности (энтропии), происходящим в результате получения сообщения, т. е.:

$$I = H_1 - H_2, \tag{13.26}$$

где I — количество информации, содержащейся в сообщении; H_1 — исходная (априорная) неопределенность; H_2 — остаточная (апостериорная) неопределенность.

Очевидно, что $H_1 > H_2$, поэтому согласно (13.26) I>0. Так, например, до тех пор, пока не были изучены статистические свойства письменных текстов, можно было априорно считать, что вероятность появления всех букв одинакова и составляет для 32- буквенного (телеграфного) алфавита $p_i = {}^1/_{32}$.

В этом случае согласно формулы энтропии:

$$H_1 = \sum_{i=1}^{i=32} \frac{1}{32} \log_2 \frac{1}{32} = 5 \text{ бит на букву.}$$

Затем путем статистического анализа множества русских текстов были установлены фактические значения вероятностей появления отдельных букв:

$$p_А = 7{,}5\%;\ p_Б = 1{,}7\%;\ \ldots p_Ю = 0{,}7\%;\ p_Я = 2{,}2\%.$$

Подстановка фактических значений вероятностей в формулу энтропии дала величину энтропии:

$$H_2 \approx 4 \text{ бит на 1 букву.}$$

В соответствии с (13.26):

$$I = H_1 — H_2 = 1 \text{ бит на 1 букву.}$$

Полученная величина I — это усредненное количество информации о статистических свойствах русских письменных текстов, которое было приобретено в результате установления фактической вероятности появления отдельных букв. Теория информации квалифицирует эту информацию как «избыточную». Названием «избыточность» подчеркивается то обстоятельство, что знание статистических характеристик передаваемых сообщений позволяет существенно разгрузить предназначенный для передачи канал связи.

Действительно, до тех пор, пока не были установлены реальные вероятности букв письменных сообщений, приходилось строить канал связи из расчета 5 бит на каждую букву. Учет реальных вероятностей уменьшает неопределенность в среднем на 1 бит на букву. На каждую 1000 переданных букв экономия составляет около 1000 бит (более подробный анализ понятия «избыточность» и способов устранения избыточности будет дан ниже).

Соотношение (13.26) распространяется на все случаи статистических исследований систем. Пока неизвестны статистические свойства исследуемой системы, наблюдатель вынужден в явной или завуалированной форме предполагать априорно равную вероятность тех или иных событий, качеств и свойств. В результате теоретических или экспериментальных исследований наблюдатель выявляет реальные значения вероятностей и тем самым уменьшает неопределенность первоначальных представлений от H_1 до H_2, т. е. в соответствии с (13.26) приобретет информацию I, равную разности H_1 — H_2. Обобщая сказанное и принимая во внимание соотношение (13.26), мы приходим к выводу, что количество информации, приобретаемое получателем сведений (сообщений), характеризует полностью или частично устраненную (снятую) неопределенность его представлений.

Рассмотренное свойство энтропии, или *правило сложения энтропии,* хорошо согласуется со смыслом энтропии как меры неопределенности. Действительно, неопределенность сообщения AB должна быть больше неопределенности отдельных сообщений A и B. Правило сложения энтропии распространяется и на n событий при $n > 2$. Доказательство этого положения аналогично приведенному выше.

Свойство 9. Энтропия характеризует среднюю неопределенность выбора одного состояния из ансамбля. При ее определении используют только вероятности состояний, полностью игнорируя их содержательную сторону. Поэтому энтропия не может служить средством решения любых задач, связанных с неопределенностью. Например, при использовании этой меры для оценки неопределенности действия лекарства, приводящего к полному выздоровлению больных в 90 % случаев и улучшению самочувствия и остальных 10 % случаев, она получится такой же, как и у лекарства, вызывающего в 90 % случаев смерть, а в 10 % — ухудшение состояния больных.

Свойство 10. Энтропия как мера неопределенности согласуется с экспериментальными данными, полученными при изучении психологических реакций человека, в частности реакции выбора. Установлено, что время безошибочной реакции на последовательность беспорядочно чередующихся равновероятных раздражителей (например, загорающихся лампочек) растет с увеличением их числа так же, как энтропия. Это время характеризует неопределенность выбора одного раздражителя.

Замена равновероятных раздражителей неравновероятными приводит к снижению среднего времени реакции ровно настолько, насколько уменьшается энтропия.

Пример 3. Заданы ансамбли U и V двух дискретных случайных величин U' и V':

$$U = \begin{vmatrix} 0,5 & 0,7 & 0,9 & 0,3 \\ 0,25 & 0,25 & 0,25 & 0,25 \end{vmatrix},$$

$$V = \begin{vmatrix} 5 & 10 & 15 & 8 \\ 0,25 & 0,25 & 0,25 & 0,25 \end{vmatrix},$$

Сравнить их энтропии.

Так как энтропия не зависит от конкретных значений случайной величины, а вероятности их появления у обеих величин одинаковы, то

$$H(U) = H(V) = \log 4 = 2 \text{ дв. ед.}$$

Выводы: *1. Если известно, что данное событие наверняка произойдет или не произойдет, то его энтропия минимальна и равна нулю.*

2. Если событие с равной вероятностью может произойти либо не произойти, то его энтропия максимальна.

3. Энтропия — величина вещественная и положительная. Для любого количества символов энтропия достигает максимума при равной вероятности появления их в сообщении.

4. Энтропия источника сообщений всегда вычисляется относительно первичного алфавита и представляет собой среднее количество информации на один качественный признак первичного алфавита.

5. Энтропия приемника сообщений представляет неопределенность появления на входе приемника декодированных сообщений, вырабатываемых источником. Энтропия приемника вычисляется относительно того же количества качественных признаков, что и энтропия источника.

6. Не следует путать энтропию источника и приемника с энтропией двух взаимонезависимых сообщений. При отсутствии помех энтропия источника равна энтропии приемника, а энтропия сообщения, состоящего из двух взаимонезависимых сообщений равна сумме их энтропии. Так же, как и энтропия n взаимонезависимых систем равна сумме энтропии каждой из n систем.

13.5. Условная энтропия

Если состояния элементов системы не зависят друг от друга, если состояние одной системы не зависит от состояния другой системы, то неопределенность того, что некоторый элемент системы (или некоторая система) будет находиться в одном из k возможных состояний полностью определялась бы вероятностными характеристиками отдельных элементов системы, либо вероятностными характеристиками состояний самих систем. В этом случае удельное количество информации на состояние элемента системы или на один символ сообщений называлось средней энтропией, а при ее вычислении использовалось выражение (13.20). При этом подразумевалось, что символы сообщения взаимонезависимы, т. е. с приходом одного символа распределение вероятностей последующих символов не изменяется. Так может быть, например, при передаче букв бесконечного алфавита, вынимаемых из кассы, либо при передаче букв конечного алфавита, но с обязательным условием, что после передачи каждой буквы она опять будет возвращена в кассу.

На практике чаще всего встречаются взаимозависимые символы и сообщения. Если передавать не просто отдельные буквы алфавита, а смысловые сообщения, то можно убедиться, что существует взаимозависимость передаваемых символов. Одни буквы встречаются чаще, другие реже, одни буквы и слова часто следуют за другими, другие редко и т. д. Например, в английском языке наиболее часто встречается буква *е,* во французском языке после буквы *q* почти всегда следует буква *и,* если *q,* естественно, не стоит в конце слова; в газетных сообщениях после слова «великолепный» чаще всего следует слово «менеджер»; появление в сообщении слов «великолепный менеджер», дает, в свою очередь, информацию о характере сообщения, например, что это сообщение ближе к бизнесовой, чем к театральной жизни и т. д.

Что касается взаимодействующих систем, то обычно состояние одной из них влияет на состояние другой, как состояние моря и скорость ветра влияют на положение корабля. В таких случаях энтропия не может быть определена только на основании безусловных вероятностей.

При подсчете среднего количества информации на символ сообщения взаимозависимость учитывают через условные вероятности свершения одних событий относительно других, а полученную при этом энтропию называют *условной энтропией.*

Рассмотрим текстовое сообщение. Вероятность появления той или иной буквы определяется тем, насколько часто встречается данная буква по сравнению с другими буквами на данное число знаков. Тот факт, что мы заранее знаем, что та или иная буква встречается чаще (или реже), чем другие буквы, уменьшает неопределенность появления этих букв и общую энтропию, характеризующую неопределенность появления букв алфавита (по сравнению с равновероятным алфавитом из того же числа качественных признаков). Так для русского алфавита *А* {—, *а, б, в, ...,я*}.

$$H(A)_{\max} = H(A)_0 = \log_2 32 = 5 \text{ бит}; \quad H(A)_1 = -\sum p_i \log_2 p_i =$$
$$= -[p(—)\log_2 p(—) + p(\text{а})\log_2 p(\text{а}) + \cdots + p(\text{я})\log_2 p(\text{я})] =$$
$$= 4{,}358 \text{ бит/букву},$$

где (—) означает пробел.

Если известны вероятности проявления одной буквы после другой (двубуквенные сочетания), то энтропия, характеризующая неопределенность появления букв алфавита, еще меньше и определяется как *условная энтропия первого порядка.* Вычисление условной энтропии будет рассмотрено ниже, а пока заметим, что неопределенность появления какой-либо буквы алфавита после данной описывается при помощи частной, условной энтропии, которая в случае энтропии первого порядка при $m_1 = 32$ имеет вид:

$$H(A/a_x) =$$
$$= -[p(a_i/a_1)\log p(a_i/a_1) +$$
$$+ p(a_i/a_2)\log p(a_i/a_2) + \cdots$$
$$\cdots + p(a_i/a_{32})\log p(a_{32}),$$

где $H(A/a_x)$ — энтропия появления буквы русского алфавита при условии, что нам известно, что ей предшествовала одна из букв $a_1 \div a_{32}$.

Например,

$$H(e/a_x) = -[p(—\text{е}) \log p(—\text{е}) + p(\text{ае}) \log p(\text{ае}) + \cdots$$
$$\cdots + p(\text{яе}) \log p(\text{яе})],$$

где p (—е); p(ае); ...; p (яе) вероятности сочетания «е» с буквами русского алфавита, включая пробел.

Если для каждой буквы алфавита известны трехбуквенные сочетания, то энтропия определяется как *условная энтропия второго порядка.* Частная условная энтропия второго порядка имеет вид:

$$H(A/a_x, a_y) = -[p(a_i/a_1, a_1) \log p(a_i/a_1, a_1) +$$
$$+ p(a_i/a_1, a_2) \log p(a_i/a_1, a_2) + \cdots + p(a_i/a_1, a_{32}) \log p(a_i/a_1, a_{32}) + \cdots$$
$$\cdots + p(a_i/a_{32}, a_{32}) \log p(a_i/a_{32}, a_{32})],$$

где $H(A/a_x, a_y)$ — энтропия появления буквы русского алфавита при условии, что ей предшествовало одно из возможных двубуквенных сочетаний, например, *аае, абе, аве, ..., аяе, ..., аяе.* Таким образом определяется вероятность трехбуквенных сочетаний.

Если для каждой буквы алфавита известны $N + 1$ буквенные сочетания, то энтропия определяется как *условная энтропия N-го порядка.*

Естественно, чем выше порядок энтропии, тем она меньше (для одного и того же алфавита). Так для русского алфавита $H_0 = 5$; $H_1 = 4{,}358$; $H_2 = 3{,}52$; $H_3 = 3{,}0$.

В случае взаимодействия N систем порядок условной энтропии определяется числом систем, взаимодействующих с данной.

Понятие условной энтропии широко используется для определения информационных потерь при передаче информации.

Рассмотрим процесс передачи сигналов по каналу связи с помехами и используем его для уяснения механизма вычисления условной энтропии.

Если элементы источника сообщений принимают состояния $a_1, a_2, ..., a_i, ..., a_n$ с вероятностями соответственно $p(a_1)$, $p(a_2)$, ..., $p(a_i)$, ..., $p(a_n)$, а элементы адресата — состояния $b_1, b_2, ..., b_j ..., b_n$ с вероятностями соответственно $p(b_1)$, $p(b_2)$, ..., $p(b_j)$, ..., $p(b_n)$, то понятие условной энтропии $H(B/a_i)$ выражает неопределенность того, что, отправив a_i, мы получим b_j, а понятие $H(A/b_j)$ — неуверенность, которая остается после получения b_j в том, что было отправлено именно a_i. Графически это может быть представлено следующим образом (рис. 13.5).

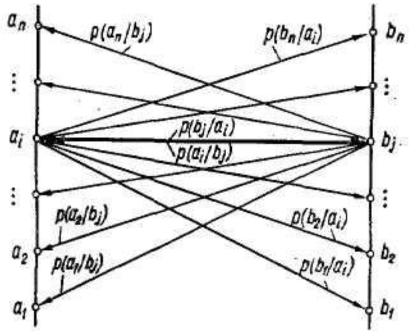


Рис. 13.5. Иллюстрация неопределенности получение сигнала b_j при передаче a_i

Посылаем сигналы a_i. Если в канале связи присутствуют помехи, то с различной степенью вероятности может быть принят любой из сигналов a_i и, наоборот, принятый сигнал b_j может появиться в результате отправления любого из сигналов a_i. Если в канале связи помехи отсутствуют, то всегда посланному символу a_1 соответствует принятый символ b_1, a_2 — b_2, ... , a_m — b_m. При этом энтропия источника $H(A)$ равна энтропии приемника $H(B)$. Если в канале связи присутствуют помехи, то они уничтожают или искажают часть передаваемой информации.

Информационные потери полностью описываются через частную и общую условную энтропию. Вычисление частных и общей условной энтропии удобно производить при помощи *канальных матриц.* Термин «канальная матрица», означающий: матрица статистически описывающая данный канал связи, применяется для краткости. Если канал связи описывается со стороны источника сообщений (т. е. известен посланный сигнал), то вероятность того, что при передаче сигнала a_i по каналу связи с помехами мы получим сигнал b_j, обозначается как условная вероятность $p\ (b_j/a_i)$, а канальная матрица имеет вид

A	B					
	b_1	b_2	$\dots$	b_j	$\dots$	b_m
a_1	$\mathbf{p(b_1/a_1)}$,	$p(b_2/a_1)$,	$\dots$,	$p(b_j/a_1)$,	$\dots$,	$p(b_m/a_1)$
a_2	$p(b_1/a_2)$,	$\mathbf{p(b_2/a_2)}$,	$\dots$,	$p(b_j/a_2)$,	$\dots$,	$p(b_m/a_2)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_i	$p(b_1/a_i)$,	$p(b_2/a_i)$,	$\dots$,	$\mathbf{p(b_j/a_i)}$,	$\dots$,	$p(b_m/a_i)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_m	$p(b_1/a_m)$,	$p(b_2/a_m)$,	$\dots$,	$p(b_j/a_m)$,	$\dots$,	$\mathbf{p(b_m/a_m)}$

(13.27)

Вероятности, которые расположены по диагонали (выделенные полужирным шрифтом), определяют вероятности правильного приема, остальные — ложного. Значения цифр, заполняющих колонки канальной матрицы, обычно уменьшаются по мере удаления от главной диагонали и при полном отсутствии помех все, кроме цифр, расположенных на главной диагонали, равны нулю.

Прохождение данного вида сигнала со стороны источника сообщений в данном канале связи описывается распределением условных вероятностей вида $p\ (b_j/a_i)$, сумма вероятностей которого всегда должна быть равна единице. Например, для сигнала a_1

$$p(b_1/a_1)+p(b_2/a_1)+\cdots+p(b_j/a_1)+\cdots+p(b_m/a_1)=1. \qquad (13.28)$$

Потери информации, приходящиеся на долю сигнала a_i, описываются при помощи *частной условной энтропии.* Например, для сигнала a_1

$$H(B/a_1)=-\sum_{i=1}^{m}p(b_j/a_1)\log p(b_j/a_1). \qquad (13.29)$$

Суммирование производится по *j*, так как *i*-е состояние (в данном случае первое) остается постоянным.

Потери при передаче всех сигналов по данному каналу связи описываются при похмощи *общей условной энтропии.* Для ее вычисления следует просуммировать все частные условные энтропии, т. е. произвести двойное суммирование по *i* и по *j*. При этом, в случае равновероятных появлений сигналов на выходе источника сообщений

$$H(B/A)=-\frac{1}{m}\sum_{j}\sum_{i}p(b_j/a_i)\log p(b_j/a_i), \qquad (13.30)$$

делим на *m,* так как энтропия есть неопределенность на один символ. В случае неравновероятного появления символов источника сообщений следует учесть вероятность появления каждого символа, умножив на нее соответствующую частную условную энтропию. При этом общая условная энтропия

$$H(B/A) = -\sum_i \sum_j p(a_i)\, p(b_j/a_i) \log p(b_j/a_i). \qquad (13.31)$$

Если исследовать канал связи со стороны приемника сообщений (то есть известен принятый сигнал), то с получением сигнала b_j предполагаем, что был послан какой-то из сигналов $a_1, a_2, ..., a_i, ... a_m$. При этом канальная матрица будет иметь вид

A	B						
	b_1	b_2		b_j		b_m	
a_1	$\mathbf{p(a_1/b_1)}$,	$p(a_1/b_2)$,	...,	$p(a_1/b_j)$.	...,	$p(a_1/b_m)$	
a_2	$p(a_2/b_1)$,	$\mathbf{p(a_2/b_2)}$,	...,	$p(a_2/b_j)$,	...,	$p(a_2/b_m)$	
...	...	...	...	...	...	...	
a_i	$p(a_i/b_1)$,	$p(a_i/b_2)$,	...,	$\mathbf{p(a_i/b_j)}$,	...,	$p(a_i/b_m)$	
...	...	...	...	...	...	...	
a_m	$p(a_m/b_1)$,	$p(a_m/b_2)$,	...,	$p(a_m/b_j)$,	...,	$\mathbf{p(a_m/b_m)}$	(13.32)

В этом случае единице должны равняться суммы условных вероятностей не по строкам, а по столбцам канальной матрицы

$$p(a_1/b_j) + p(a_2/b_j) + \cdots + p(a_i/b_j) + \cdots + p(a_m/b_j) = 1.$$

Частная условная энтропия

$$H(A/b_j) = -\sum_{i=1}^{m} p(a_i/b_j) \log p(a_i/b_j). \qquad (13.33)$$

А общая условная энтропия

$$H(A/B) = -\sum_j \sum_i p(b_j)\, p(a_i/b_j) \log p(a_i/b_j). \qquad (13.34)$$

Пример. Сообщения на выходе источника создаются путем комбинирования частот f_1, f_2, f_3. Статистические испытания канала связи при прохождении этих частот дали результаты, представленные следующей канальной матрицей:

$$p(b/a) = \begin{Vmatrix} 0{,}98 & 0{,}01 & 0{,}01 \\ 0{,}1 & 0{,}75 & 0{,}15 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{Vmatrix}.$$

Определить неопределенность прохождения частоты f_2 по данному каналу связи. Определить общую условную энтропию данного канала связи, если вероятности появления частот на выходе источника сообщений $p(f_1) = 0{,}7$; $p(f_2) = 0{,}2$; $p(f_3) = 0{,}1$.

Решение: 1) Частная условная энтропия

$$H(B/f_2) = -(0{,}1\log_2 0{,}1 + 0{,}75\log_2 0{,}75 + 0{,}15\log_2 0{,}15) = 0{,}3322 + 0{,}3113 + 0{,}4105 = 1{,}054 \text{ бит/символ.}$$

2) Общая условная энтропия

$$H(B/A) = -\sum_i \sum_j p(a_i)\, p(b_j/a_i) \log_2 p(b_j/a_i) =$$

$$= -[0{,}7(0{,}98\log_2 0{,}98 + 2\cdot 0{,}01\log_2 0{,}01) + 0{,}2(0{,}1\log_2 0{,}1 + 0{,}75\log_2 0{,}75 + 0{,}15\log_2 0{,}15) + 0{,}1(0{,}2\log_2 0{,}2 + 0{,}3\log_2 0{,}3 + 0{,}5\log_2 0{,}5)] =$$

$$= 0{,}7(0{,}0285 + 2\cdot 0{,}0664) + 0{,}2\cdot 1{,}054 + 0{,}1(0{,}4644 + 0{,}5211 + 0{,}5) =$$

$$= 0{,}46526 \text{ бит/символ.}$$

Примечание. Общая условная энтропия получилась меньше частной потому, что при вычислении $H(B/f_2)$ не учитывалась вероятность появления сигнала f_2 на выходе источника сообщений.

Если в канале связи помехи отсутствуют, то все элементы канальной матрицы, кроме расположенных на главной диагонали, равны нулю. Это говорит о том, что при передаче сигнала a_1 мы наверняка получим b_1, при передаче a_2 — b_2, ..., a_m — b_m. Вероятность получения правильного сигнала станет безусловной, а условная энтропия будет равна нулю.

Своего максимума условная энтропия достигает в случае, когда при передаче символа a_i может быть с равной вероятностью получен любой из сигналов $b_1, b_2, \ldots, b_m$. Строго говоря, энтропия шумящего канала может быть больше энтропии источника сообщений (например, если известны вероятностные характеристики появления букв на выходе источника сообщений, то его энтропия будет меньше энтропии канала связи для передачи количества качественных признаков, равного числу букв на выходе источника сообщений, если уровень помех в канале связи настолько высок, что с равной вероятностью каждый качественный признак может перейти в любой другой).

При вычислении условной энтропии механических, экономических и других взаимосвязанных систем могут быть использованы матрицы, аналогичные вышеописанным. Однако следует иметь в виду, что матрицы, описывающие каналы связи, всегда квадратные (так как сколько качественных признаков передаем, столько предполагаем принять, хотя в некоторых случаях вероятность прохождения отдельных сигналов может быть равна нулю). Но при матричном описании механических или других взаимосвязанных систем соотношение столбцов и строк матрицы может быть произвольным, так как один или группа элементов исследуемой системы может взаимодействовать с произвольным количеством других элементов. По этим же причинам пределы суммирования по i и по j могут быть разные.

Предположим, имеются две системы: A с возможными состояниями $i = 1,2,3, \ldots m_1$ и B с состояниями $j = 1, 2, 3, \ldots, m_2$. Энтропия системы B при условии, что система A находится в состоянии a_i:

$$H(B/a_i) = -\sum_{j=1}^{m_2} p(b_j/a_i) \log p(b_j/a_i) = M[-\log p(b_j/a_i)], \tag{13.35}$$

где $p(b_j/a_i)$ есть вероятность нахождения системы B в состоянии b_j, если система A находится в состоянии a_i, $M[—\log p(b_j/a_i)]$ — математическое ожидание величины $[—\log p(b_j/a_i)]$.

Общая условная энтропия системы B относительно системы A

$$H(B/A) = \sum_{i=1}^{m_1} p(a_i) H(B/a_i) = -\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} p(a_i) p(b_j/a_i) \log p(b_j/a_i).$$

Равенство m_1 и m_2 не обязательно.

Пример. Определить общую условную энтропию двух взаимодействующих систем X и Y, если известны вероятности состояний элементов системы Y : $p(y_1) = 0{,}2$; $p(y_2) = 0{,}3$; $p(y_3)$ - $0{,}4$; $p(y_4) = 0{,}1$, а матрица, описывающая взаимозависимость системы X и Y имеет вид:

$$p(x/y) = \begin{Vmatrix} 0{,}1 & 0{,}5 & 0{,}3 & 0{,}1 \\ 0{,}9 & 0{,}5 & 0{,}7 & 0{,}9 \end{Vmatrix}$$

Решение: Общая условная энтропия

$$H(X/Y) = -\sum_{j=1}^{4} \sum_{i=1}^{2} p(y_j) p(x_i/y_j) \log_2 p(x_i/y_j) = -[0{,}2(0{,}1 \log_2 0{,}1 + 0{,}9 \log_2 0{,}9) + \\ + 0{,}3(0{,}5 \log_2 0{,}5 + 0{,}5 \log_2 0{,}5) + 0{,}4(0{,}3 \log_2 0{,}3 + 0{,}7 \log_2 0{,}7) + \\ + 0{,}1(0{,}1 \log_2 0{,}1 + 0{,}9 \log_2 0{,}9)] = 0{,}09378 + 0{,}3 + 0{,}35262 + 0{,}04689 = \\ = 0{,}79329 \text{ бит/состояние.}$$

Общая условная энтропия системы B относительно системы A характеризует количество информации, содержащейся в любом символе абстрактного алфавита, через который мы представляем состояния элементов исследуемых систем.

Общую условную энтропию определяют путем усреднения по всем символам, т. е. по всем состояниям a_1 с учетом вероятности появления каждого из них. Она равна сумме вероятностей появления символов алфавита на неопределенность, которая остается после того, как адресат принял сигнал:

$$H(B/A) = -\sum_i p(a_i) H(b_j/a_i) = -\sum_i \sum_j p(a_i) p(b_j/a_i) \log p(b_j/a_i). \tag{13.36}$$

Выражение (13.36) является общим для определения количества информации на один символ сообщения в случае взаимозависимых и неравновероятных символов. Поскольку $p\ (a_i)\ p\ (b_j/a_i)$ представляет собой вероятность совместного появления двух событий $p\ (a_i,\ b_j)$, то (13.36) можно записать следующим образом:

$$H(B/A) = -\sum_i \sum_j p(a_i,\ b_j) \log p(b_j/a_i). \tag{13.37}$$

Итак, в случае взаимозависимых и равновероятных символов в сообщении, т. е. при

$$p(a_i) = p(b_j) = \frac{1}{m},$$

$$H(A/B) = -\frac{1}{m} \sum_j p(a_i/b_j) \log p(a_i/b_j). \tag{13.38}$$

В случае неравновероятных, но независимых символов, т. е. при

$$p\ (a_i/b_j) = p\ (a_i)$$

$$H(A) = -\sum p\ (a_i) \log p\ (a_i). \tag{13.39}$$

И наконец, в случае равновероятных и независимых символов

$$p(a_i) = p(b_j) = \frac{1}{m};$$

$$H = -\log \frac{1}{m} = \log m, \tag{13.40}$$

т. е. приходим к формуле Хартли. Это предельное количество информации, составленное из равновероятных и независимых символов и позволяющее максимально использовать символы алфавита, Шеннон назвал *максимальной энтропией,* тем самым установив предел информационной нагрузки сообщения из ограниченного числа символов.

Уяснению соотношений между рассмотренными энтропиями дискретных источников информации (ансамблей) способствует их графическое отображение (рис. 13.6).

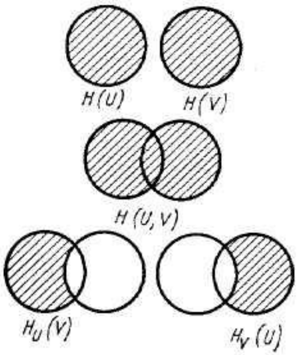


Рис. 13.6

Пример. Известны энтропии двух зависимых источников: $H(U)$ = 5 дв. ед., $H(V)$ = 10 дв. ед. Определить, в каких пределах будет изменяться условная энтропия $H_U(V)$ при изменении $H_V(U)$ в максимально возможных пределах.

При решении удобно использовать графическое отображение связи между энтропиями. Из рис. 13.7 видим, что максимального значения $H_U(V)$ достигает при отсутствии взаимосвязи и будет равно $H(V)$, т.е. 10 дв. ед.

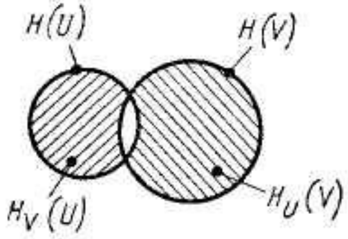


Рис. 13.7

По мере увеличения взаимосвязи $H_U(V)$ будет уменьшаться до значения $H(V) — H(U)$ = 5 дв. ед. При этом $H_U(V)$ = 0.

Выражения (13.36) — (13.40) могут быть использованы для подсчета количества информации как для отдельных элементов, так и для сообщения в целом.

При небольшом количестве символов в алфавите вероятности p_i легко задать. Для сообщений, составленных из k неравновероятных взаимозависимых символов, при достаточно большом числе k (т. е. при длинных сообщениях) можно использовать формулу (13.39), так как при этом отдельные символы становятся практически независимыми. В противном случае определить количество информации невозможно без таблицы условных вероятностей.

Если заранее известны статистические свойства источника сообщений, в котором учтены распределения вероятностей отдельных символов (например, статистические свойства различных буквенных алфавитов хорошо изучены и вероятности появления отдельных букв могут быть заданы априори) и взаимосвязи между ними (например, появление буквы h после t в английском языке), а также характер распределения помех в канале связи (например, свойства белого шума хорошо изучены, известно, что у него нормальный закон распределения и даже найдено выражение для энтропии суммы сигнала и шума, то для вычисления количества информации выписывают значения вероятностей отдельных символов и исходных условных вероятностей и подставляют их в (13.36). В противном случае соответствующие вероятности определяют опытным путем.

Напомним, как это делается. Пусть при передаче n сообщений символ A появился m раз, т. е. $p(A) = m/n$; символ B — l раз, т. е. $p(A + B) = m/n + l/n$. Далее, пусть при передаче n сообщений символ B появился l, а символ A вместе с символом B — k раз. Вероятность появления символа B $p(B) = l/n$, вероятность совместного появления символов A и B $p(AB) = k/n$, условная вероятность появления символа A относительно символа B

$$p(A/B) = \frac{p(AB)}{p(B)} = \frac{k}{l}. \qquad (13.41)$$

Если известна условная вероятность, то можно легко определить вероятность совместного появления символов A и B, используя (13.41):

$$p(AB) = p(B)\,p(A/B) = p(A)\,p(B/A).$$

Если известно, что среди n сообщений был получен символ A и что в сообщениях $E_1, E_2, \ldots, E_n$ присутствует символ A, и мы хотим определить вероятность получения сообщения $E_1, E_2, \ldots, E_n$, зная о получении символа A, то пользуемся формулой Бейеса:

$$p(E_k/A) = \frac{p(E_k)\,p(A/E_k)}{\sum\limits_{i=1}^{n} p(E_i)\,p(A/E_i)},$$

где $p(E_k/A)$ — условная вероятность появления сообщений относительно появления символа A:

$$\sum_{i}^{n} p(E_i)\,p(A/E_i) = p(A) \text{ — полная вероятность события.}$$

В заключение рассмотрим несколько основных свойствах условной энтропии.

Свойство 1. *Если ансамбли сообщений A и B взаимонезависимы, то условная энтропия A относительно B равна безусловной энтропии A:*

$$H(A/B) = H(A); \quad H(B/A) = H(B).$$

Доказательство. Если сообщения A и B взаимонезависимы, то условные вероятности отдельных символов равны безусловным:

$$p(b_j/a_i) = p(b_j). \qquad (13.42)$$

Так как $p(a_i)\,p(b_j/a_i) = p(a_i, b_j)$, то выражение (13.36) представим в виде:

$$H(B/A) = -\sum_{i}\sum_{j} p(a_i, b_j) \log p(b_j/a_i),$$

подставив в полученное выражение (13.42), получим

$$H(B/A) = -\sum_{i}\sum_{j} p(a_i, b_j) \log p(b_j) = -\sum_{j} p(b_j) \log p(b_j) = H(B),$$

так как

$$\sum_i p\ (a_i,\ \ b_j) = p\ (b_j).$$

Свойство 2. *Если ансамбли сообщений А и В настолько жестко статистически связаны, что появление одного из них непременно подразумевает появление другого, то их условные энтропии равны нулю:*

$$H\ (A/B) = H\ (B/A) = 0.$$

Доказательство: Воспользуемся свойствами вероятностей, согласно которым при полной статистической зависимости $p(b_j/a_i)=1$ слагаемые $p\ (b_j/a_i)\ \log\ p\ (b_j/a_i)$ выражения (13.36) также равны нулю. Если нулю равны отдельные слагаемые, то и сумма равна нулю, откуда

$$H\ (B/A) = H\ (A/B) = 0.$$

Выводы: *1. Энтропия сообщения, составленного с учетом неравновероятности символов меньше, чем энтропия сообщения, составленного из равновероятных символов.*

2. Энтропия сообщения, составленного с учетом взаимозависимости символов, меньше, чем энтропия того же сообщения, составленного из взаимонсзависимых символов.

3. Максимальную энтропию имеют сообщения, составленные из равновероятных и независимых символов, т. е. те, у которых условная энтропия равна нулю, а вероятность появления символов алфавита $p = 1/m$.

4. Информационные характеристики канала связи можно использовать для улучшения точностных характеристик, передаваемых сообщений путем проведения двойного упорядочивания: сначала определить потери на кодовое слово, упорядочить кодовые слова по убыванию потерь, упорядочить первичный алфавит по убыванию вероятностей, присвоить упорядоченную последовательность кодовых слов упорядоченной последовательности букв первичного алфавита.

13.6. Энтропия объединения

Энтропия объединения используется для вычисления энтропии совместного появления статистически зависимых сообщений либо энтропии взаимосвязанных систем.

Например, при передаче по каналу связи с шумами цифры 5, из 100 раз цифра 5 была принята 90 раз, цифра 6 — восемь раз, цифра 4 — два раза. Неопределенность возникновения комбинаций вида 5— 4; 5—5; 5—6 при передаче цифры 5 может быть описана при помощи энтропии объединения *H (A, B)*.

Другой пример. Движение автомобиля на одной из передач возможно лишь при отжатом сцеплении, направление движения зависит от положения руля, скорость — от положения акселератора и т. д. Таким образом направление, скорость и сам факт движения несет в себе информацию о состоянии и характере взаимодействия некоторых элементов системы «автомобиль». Взаимозависимость элементов системы «автомобиль» может быть описана при помощи энтропии объединения.

Восприятие понятия «энтропия объединения» (иногда употребляют термин «взаимная энтропия») облегчается, если рассуждения сводятся к некоторому условному каналу связи. На примере передачи информации по каналу связи также удобнее проследить взаимосвязь понятий условной энтропии *H (B/A)* и взаимной энтропии *H (A, B).*

Итак, пусть $(a_1, a_2, ..., a_i.....a_n)$ есть выборочное пространство *A,*

характеризующее источник сообщений, а $(b_1, b_2, ..., b_j, ..., b_m)$ есть выборочное пространство *B,* характеризующее приемник сообщений. При этом *a* есть сигнал на входе шумящего канала, а *b* — сигнал на его выходе. Взаимосвязь переданных и принятых сигналов описывается вероятностями совместных событий вида *p (a, b),* а взаимосвязь выборочных пространств *A* и *B* описывается матрицей вида:

$$p(a_i, b_j) = \begin{Vmatrix} p(a_1, b_1) & p(a_1, b_2) & \dots & p(a_1, b_m) \\ p(a_2, b_1) & p(a_2, b_2) & \dots & p(a_2, b_m) \\ \dots & \dots & \dots & \dots \\ p(a_n, b_1) & p(a_n, b_2) & \dots & p(b_n, b_m) \end{Vmatrix} \tag{13.43}$$

Если матрица описывает канал связи, то число строк матрицы равно числу столбцов *m — n* и пределы суммирования по *i* и по *j* одинаковы. При описании взаимодействия систем равенство *m* и *n* необязательно (см. предпоследний пример п.13.5).

Независимо от равенства или неравенства числа строк числу столбцов матрица объединения (13.43) обладает свойством:

$$\sum_i p(a_i, b_j) = p(b_j); \tag{13.44}$$

$$\sum_j p(a_i, b_j) = p(a_i). \tag{13.45}$$

В свою очередь,

$$\sum_i p(a_i) = \sum_j p(b_j) = 1,$$

т. е.

$$\sum_i \sum_j p(a_i, b_j) = 1.$$

Это свойство позволяет вычислять энтропию источника и приемника сообщений непосредственно по матрице объединения (13.43):

$$H(A) = -\sum_i \sum_j p(a_i, b_j) \log \sum_j p(a_i, b_j); \tag{13.46}$$

$$H(B) = -\sum_i \sum_j p(a_i, b_j) \log \sum_i p(a_i, b_j). \tag{13.47}$$

До знака логарифма суммирование производится по *i* и *j*, так как для нахождения безусловных вероятностей необходимо производить суммирование по одной координате (имеется в виду матричное представление вероятностей), а для нахождения соответствующей энтропии суммирование производится по другой координате.

Условные вероятности при помощи матрицы объединения находятся следующим образом:

$$p(a_i/b_j) = \frac{p(a_i, b_j)}{\sum_i p(a_i, b_j)} = \frac{p(a_i, b_j)}{p(b_j)}; \tag{13.48}$$

$$p(b_j/a_i) = \frac{p(a_i, b_j)}{\sum_j p(a_i, b_j)} = \frac{p(a_i, b_j)}{p(a_i)}. \tag{13.49}$$

Энтропия объединения ансамблей *A* и *B* при помощи матрицы (13.43) вычисляется путем последовательного суммирования по строкам или по столбцам всех вероятностей вида *p (a, b)*, умноженных на логарифм этих же вероятностей

$$H(A, B) = -\sum_i \sum_j p(a_i, b_j) \log_2 p(a_i, b_j) \text{ бит/два символа}. \tag{13.50}$$

Размерность «бит/два символа» объясняется тем, что взаимная энтропия представляет собой неопределенность возникновения пары символов, т. е. неопределенность на два символа. В случае отсутствия взаимозависимости между символами выражение (13.50) принимает вид выражения (13.51), соответственно и размерность будет «бит/символ».

Исследуем выражение (13.50).

Из теории вероятностей известно, что

$$p(a_i, b_j) = p(a_i)\, p(b_j/a_i) = p(b_j)\, p(a_i/b_j). \tag{13.51}$$

Используя (13.51), выражение (13.50) можно записать как

$$H(A,B) = -\sum_i \sum_j p(a_i)\, p(b_j/a_i) \log p(a_i)\, p(b_j/a_i) =$$
$$= -\left[\sum_i \sum_j p(a_i)\, p(b_j/a_i) \log p(a_i) + \sum_i \sum_j p(a_i)\, p(b_j/a_i) \log p(b_j/a_i)\right].$$
(13.52)

Но $p(a_i)\, p(b_j/a_i) = p(a_i, b_j)$, а согласно (13.45),

$$\sum_j p(a_i, b_j) = p(a_i),$$

тогда первое слагаемое выражения (13.52) принимает вид

$$\sum_i p(a_i) \log p(a_i) = -H(A).$$

Второе слагаемое, согласно (13.31), есть не что иное, как $H(B/A)$, что позволяет выражение (13.52) записать в виде

$$H(A,B) = H(A) + H(B/A). \qquad (13.53)$$

Энтропия объединения передаваемого ансамбля A и принимаемого ансамбля B равна сумме безусловной энтропии $H(A)$ и условной энтропии $H(B/A)$.

Последняя в данном случае представляет ту добавочную информацию, которую дает сообщение B после того, как стала известна информация, содержащаяся в сообщении A.

Таким образом, условная энтропия представляет собой неопределенность того, что при приеме b было послано a, а взаимная энтропия отражает неопределенность возникновения пары вида ab.

Так как энтропия объединения есть неопределенность относительно пары символов, сигналов, состояний, в общем случае, относительно пары элементов взаимосвязанных выборочных пространств A и B, то вид пары ab или ba не имеет значения, так как неопределенность возникновения такого сочетания — одинакова. *Энтропия объединения обладает свойством симметрии*

$$H(A, B) = H(B, A).$$

Свойство симметрии энтропии объединения хорошо иллюстрируется матрицей (13.43). Действительно, если просуммировать все элементы матрицы объединения по строкам и по столбцам по схеме рис. 13.8, а затем сложить полученные результаты, то обе суммы будут равны единице.

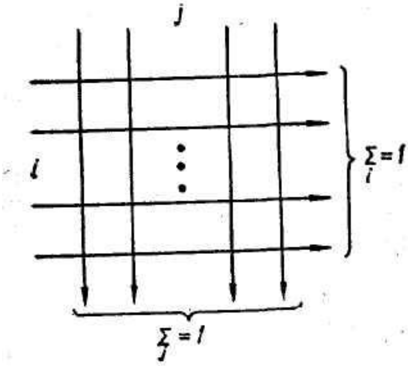


Рис. 13.8. Сумма элементов по строкам и столбцам матрицы объединения равна единице

Свойство симметрии позволяет записать соотношения между условной энтропией и энтропией объединения следующим образом:

$$H(A, B) = H(A) + H(B/A) = H(B) + H(A/B); \qquad (13.54)$$

$$H(B/A) = H(A, B) - H(A); \quad H(A/B) = H(A, B) - H(B). \qquad (13.55)$$

Если построена матрица вероятностей *p(a, b),* описывающая взаимосвязь двух произвольных выборочных пространств, в частности взаимосвязь входа и выхода шумящего канала связи, то остальные информационные характеристики могут не задаваться, так как матрица (13.43) обладает *информационной полнотой.*

Определение 1. Набор информационных характеристик является информационно полным, если с помощью этого набора можно путем алгебраических преобразований получить любую другую информационную характеристику того же канала связи.

Определение 2. Полностью описанным в информационном плане считается канал связи, для которого может быть определена любая из его информационных характеристик.

Побочными характеристиками, которые могут быть при этом получены, являются вероятность ошибки в данном канале связи и необходимая избыточность сообщений для восстановления искажаемой помехами информации.

Утверждение 1. Информационной полнотой обладает набор информационных характеристик, если он содержит: $p(a_i)$ и $p(b_j/a_i)$ или $p(b_j)$ и $p(a_i/b_j)$, или $p(a_i, b_j)$.

Доказательство: 1. При известных $p(a_i)$ и $p(b_j/a_i)$

$$H(A) = -\sum_{i=1}^{m} p(a_i)\log p(a_i); \quad H(B/a) = -\sum_{i=1}^{m} p(b_j/a_i)\log p(b_j/a_i);$$

$$H(B/A) = -\sum_{i=1}^{m}\sum_{j=1}^{m} p(a_i)\,p(b_j/a_i)\log p(b_j/a_i).$$

Матрица вероятностей $p\ (b_j/a_i)$ представлена на рис. 13.9.

$$p(b_j/a_i) = \left\| \begin{array}{cccc} p(b_1/a_1) & p(b_2/a_1) & \dots & p(b_m/a_1) \\ p(b_1/a_2) & p(b_2/a_2) & \dots & p(b_m/a_2) \\ \vdots & \vdots & & \vdots \\ p(b_1/a_m) & p(b_2/a_m) & \dots & p(b_m/a_m) \end{array} \right\| \begin{array}{l} \sum_j p(b_j/a_1)=1 \\ \sum_j p(b_j/a_2)=1 \\ \vdots \\ \sum_j p(b_j/a_m)=1 \end{array}$$

Рис. 13.9. Матрица вероятностей вида $p\ (b_j/a_i)$

Произведение $p(a_i)$ с элементами матрицы $p(b_j/a_i)$ по схеме, представленной на рис. 13.10, дает матрицу вероятностей вида $p\ (a_i, b_j)$, представленную на рис. 13.11:

$$H(B) = -\sum_j p(b_j)\log p(b_j).$$

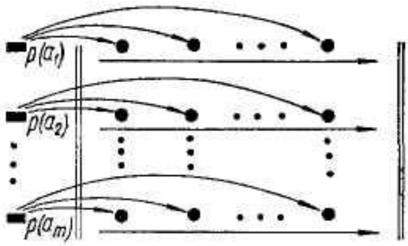


Рис. 13.10. Схема умножения элементов матрицы рис. 13.9 на $p\ (a_i)$

$$p(a_i,b_j) = \left\| \begin{array}{cccc} p(a_1,b_1) & p(a_1,b_2) & \dots & p(a_1,b_m) \\ p(a_2,b_1) & p(a_2,b_2) & \dots & p(a_2,b_m) \\ \vdots & \vdots & & \vdots \\ p(a_m,b_1) & p(a_m,b_2) & \dots & p(a_m,b_m) \end{array} \right\| \begin{array}{l} \sum_j = p(a_1) \\ \sum_j = p(a_2) \\ \vdots \\ \sum_j = p(a_m) \end{array}$$

$$\sum_i = p(b_1) \quad \sum_i = p(b_2) \quad \dots \quad \sum_i = p(b_m)$$

Рис. 13.11. Матрица вероятностей вида $p\ (a_i,\ b_j)$

Деление элементов матрицы $p\ (a_{i,}\ b_j)$ на $p\ (b_j)$ по схеме рис. 13.12 дает матрицу вида $p\ (a_i/b_j)$ (рис. 13.13):

$$H\ (A/b) = -\sum_{i=1}^{m} p\ (a_i/b_j) \log p\ (a_i/b_j);$$

$$H\ (A/B) = -\sum_{i=1}^{m} \sum_{j=1}^{m} p\ (b_j)\ p\ (a_i/b_j) \log p\ (a_i/b_j);$$

$$I\ (A,\ B) = H\ (A) - H\ (A/B) = H\ (B) - H\ (B/A) =$$
$$= H\ (A) + H\ (B) - H\ (A,\ B).$$

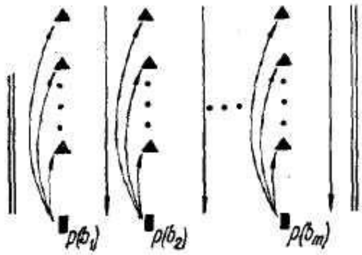


Рис. 13.12. Схема деления элементов матрицы рис. 13.11 на $p\ (b_j)$

$$p(a_i/b_j) = \left\| \begin{array}{c|c|c} \circ\, p(a_1/b_1) & \circ\, p(a_1/b_2)\ \dots & \circ\, p(a_1/b_m) \\ \circ\, p(a_2/b_1) & \circ\, p(a_2/b_2)\ \dots & \circ\, p(a_2/b_m) \\ \vdots & \vdots & \vdots \\ \circ\, p(a_m/b_1) & \circ\, p(a_m/b_2)\ \dots & \circ\, p(a_m/b_m) \end{array} \right\|$$

$$\sum_i = p(a_i/b_1) = 1 \quad \sum_i = p(a_i/b_2) = 1 \cdots \sum_i = p(a_i/b_m) = 1$$

Рис. 13.13. Матрица вероятностей вида $p\ (a_i/b_j)$

2. При известных $p\ (b_j)$ и $p\ (a_i/b_j)$

$$H\ (B) = -\sum_{j=1} p\ (b_j) \log p\ (b_j);$$

$$H\ (A/b) = -\sum_{i=1}^{m} p\ (a_i/b_j) \log p\ (a_i/b_j);$$

$$H\ (B/A) = -\sum_{j=1}^{m} \sum_{i=1}^{m} p\ (b_j)\ p\ (a_i/b_j) \log p\ (a_i/b_j).$$

Произведение $p\,(b_j)$ с элементами матрицы $p\,(a_i/b_j)$, представленной на рис. 13.13, дает матрицу вида $p\,(b_j,\ a_i)$ (рис. 13.14). Деление элементов матрицы $p\,(b_j,\ a_i)$ на $p\,(a_i\ )$ по схеме рис. 13.15 дает матрицу вида $p\,(b_j/a_i)$, тогда

$$H\,(B/a) = -\sum_{j=1}^{m} p\,(b_j/a_i)\log p\,(b_j/a_i);$$

$$H\,(B/A) = -\sum_{i=1}^{m}\sum_{j=1}^{m} p\,(a_i)\,p\,(b_j/a_i)\log p\,(b_j/a_i);$$

$$I\,(B,\ A) = H\,(A) - H\,(A/B) = H\,(B) - H\,(B/A) = \\ = H\,(A) + H\,(B) - H\,(B/A).$$

3. При известных $p\,(a_i,\ b_j)$

$$\sum_j p\,(a_i,\ b_j) = p\,(a_i);$$

$$\sum_i p\,(a_i,\ b_j) = p\,(b_j).$$

При делении матрицы вида $p\,(a_i,\ b_j)$ на $p\,(b_j)$ по схеме рис. 13.12 получаем матрицу вероятностей $p\,(a_i\ /b_j)$, а при делении на $p\,(a_i)$ по схеме рис. 13.15 получим матрицу вида $p\,(b_j/a_i)$, что позволяет найти *H (A/b), H (B/a), H (B/A), H (A/B)* и *H(A, B) = H (B, A)*.

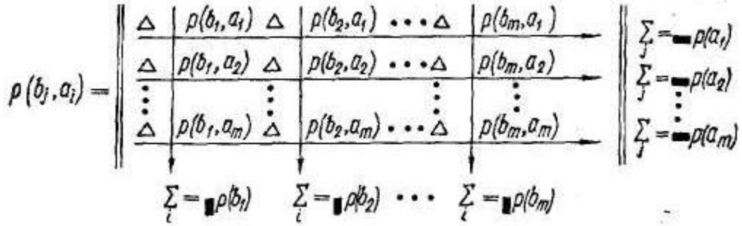


Рис. 13.14 Матрица вероятностей вида $p\,(b_j,\ a_i)$

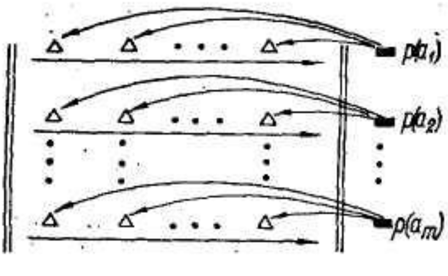


Рис. 13.15 Схема деления элементов матрицы рис. 13.14 на $p\,(a_i)$

Пример. Задана матрица вероятностей системы, объединенной в одну из двух систем *A* и *B:*

$$p(A,\ B)=\begin{Vmatrix} 0{,}3 & 0 & 0 \\ 0{,}2 & 0{,}3 & 0{,}1 \\ 0 & 0{,}1 & 0 \end{Vmatrix}$$

Определить безусловную энтропию системы *A,* безусловную энтропию системы *B,* полные условные энтропии *H (B/A)* и *H (A/B),* а также взаимную энтропию *H (A, B).*

Решение: 1) Вычисляем безусловные вероятности как суммы совместных вероятностей по строкам и столбцам исходной матрицы

$$\begin{array}{rcc} & & p(a_i) \\ p(A,\ B)= & \begin{Vmatrix} 0{,}3 & 0 & 0 \\ 0{,}2 & 0{,}3 & 0{,}1 \\ 0 & 0{,}1 & 0 \end{Vmatrix} & \begin{matrix} 0{,}3 \\ 0{,}6 \\ 0{,}1 \end{matrix} \\ \hline p(b_j) & \begin{matrix} 0{,}5 & 0{,}4 & 0{,}1 \end{matrix} & \end{array}$$

$$H(A)=-\sum_{j=1}^{m} p(a_i)\log_2 p(a_i)=-(0{,}3\log_2 0{,}3+0{,}6\log_2 0{,}6+0{,}1\log_2 0{,}1)=$$

$$=1{,}295\ \text{бит/состояние;}$$

$$H(B)=-\sum_{j=1}^{m} p(b_j)\log_2 p(b_j)=-(0{,}5\log_2 0{,}5+0{,}4\log_2 0{,}4+0{,}1\log_2 0{,}1)=$$

$$=1{,}36\ \text{бит/состояние.}$$

2) Определяем условные вероятности и составляем матрицу условных вероятностей:

$$p(a_i,\ b_j)=\frac{p(a_i,\ b_j)}{p(b_j)};\quad p(a_1/b_1)=\frac{0{,}3}{0{,}5}=0{,}6;\quad p(a_2/b_2)=\frac{0{,}3}{0{,}4}=0{,}75;$$

$$p(a_3/b_2)=\frac{0{,}1}{0{,}4}=0{,}25;\quad p(a_2/b_3)=\frac{0{,}1}{0{,}1}=1;\quad p(a_2/b_1)=p(a_1/b_2)=p(a_1/b_3)=$$

$$=p(a_3/b_3)=0.$$

$$p(a_i/b_j)=\begin{Vmatrix} 0{,}6 & 0 & 0 \\ 0{,}4 & 0{,}75 & 1 \\ 0 & 0{,}25 & 0 \end{Vmatrix}$$

$$H(A/B)=-\sum_i\sum_j p(b_j)\,p(a_i/b_j)\log_2 p(a_i/b_j)=-[0{,}5(0{,}6\log_2 0{,}6+0{,}4\log_2 0{,}4)+$$

$$+0{,}4(0{,}75\log_2 0{,}75+0{,}25\log_2 0{,}25)+(1\log_2 1)\cdot 0{,}1]\approx 0{,}324+0{,}485=$$

$$=0{,}809\ \text{бит/состояние,}$$

или

$$H(A/B)=-\sum_i\sum_j p(a_i,\ b_j)\log_2 p(b_j/a_i)=-(0{,}3\log_2 0{,}6+0{,}2\log_2 0{,}4+$$

$$+0{,}3\log_2 0{,}75+0{,}1\log_2 0{,}25)=0{,}3\cdot 0{,}736+0{,}2\cdot 1{,}321+0{,}3\cdot 0{,}415+$$

$$+0{,}1\cdot 2\approx 0{,}809\ \text{бит/состояние.}$$

3) Аналогично для *H (B/A):*

$$p(b_j/a_i) = \frac{p(a_i, b_j)}{p(a_i)}; \quad p(b_1/a_1) = \frac{0,3}{0,3} = 1;$$

$$p(b_1/a_2) = \frac{0,2}{0,6} = 0,333; \quad p(b_2/a_2) = \frac{0,3}{0,6} = 0,5;$$

$$p(b_2/a_3) = \frac{0,1}{0,1} = 1; \quad p(b_3/a_2) = \frac{0,1}{0,6} = 0,167; \quad p(a_3/b_1) = p(a_1/b_2) =$$

$$= p(a_1/b_3) = p(a_3/b_3) = 0.$$

$$p(b_j/a_i) = \begin{Vmatrix} 1 & 0 & 0 \\ 0,333 & 0,5 & 0,167 \\ 0 & 1 & 0 \end{Vmatrix}$$

$$H(B/A) = -\sum_i \sum_j p(a_i)\, p(b_j/a_i) \log_2 p(b_j/a_i) = -0,6\,(0,333 \log_2 0,333 +$$

$$+ 0,5 \log_2 0,5 + 0,167 \log_2 0,167) \approx 0,875 \text{ бит/состояние,}$$

или

$$H(B/A) = -\sum_i \sum_j p(a_i, b_j) \log_2 p(b_j/a_i) = -(0,2 \log_2 0,333 + 0,3 \log_2 0,5 +$$

$$+ 0,1 \log_2 0,167) \approx 0,2 \cdot 1,586 + 0,1 \cdot 2,582 = 0,8754 \text{ бит/состояние.}$$

4) Определяем взаимную энтропию

$$H(A, B) = -\sum_i \sum_j p(a_i, b_j) \log_2 p(a_i, b_j) = -(2 \cdot 0,3 \log_2 0,3 +$$

$$+ 2 \cdot 0,1 \log_2 0,1 + 0,2 \log_2 0,2) = 2,17 \text{ бит/два состояния.}$$

Проверка:

$$H(A, B) = H(A) + H(B/A) = 1,295 + 0,8754 = 2,169;$$

$$H(B, A) = H(B) + H(A/B) = 1,36 + 0,809 = 2,169.$$

Рассмотренный выше пример наглядно иллюстрирует соотношения между различными информационными характеристиками ансамблей, у которых наблюдается статистическая зависимость.

При отсутствии статистической зависимости между элементами ансамблей *A* и *B* или между элементами некоторой системы, объединенной в одну из двух систем *A* и *B,* условные вероятности превращаются в безусловные

$$p(b_j/a_i) = p(b_j) \quad \text{и} \quad p(a_i/b_j) = p(a_i).$$

В этом случае

$$H(A, B)=H(A) + H(B).$$

Действительно

$$H(A, B) = \mathbf{M}[-\log p(a, b)] = -\sum_{i,j} p(a_i, b_j) \log_2 p(a_i, b_j),$$

если $p(b_j/a_i) = p(b_j)$ и $p(a_i/b_j) = p(a_i)$, то

$$p(a_i, b_j) = p(a_i)\,p(b_j); \quad \log p(a_i, b_j) = \log p(a_i) + \log p(b_j);$$
$$\mathbf{M}[-\log p(a_i, b_j)] = \mathbf{M}[-\log p(a_i)] + \mathbf{M}[-\log p(b_j)];$$
$$H(A, B) = H(A) + H(B).$$

При полной статистической зависимости между элементами ансамблей A и B (например, когда результат одного события однозначно определяет информацию о другом событии)
$H(B/A) = H(A/B) = 0$, а взаимная энтропия

$$H(A, B) = H(A) = H(B). \tag{13.56}$$

В случае передачи информации по каналам связи полная статистическая зависимость между передаваемыми и принимаемыми сигналами говорит об отсутствии помех, канальная матрица приобретает вид

$$p(b_j/a_i) = \begin{Vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{Vmatrix}$$

условные вероятности правильного приема равны единице, а остальные — нулю, что превращает в нуль все частные условные энтропии.

$$H(B/a_i) = -\sum_j p(b_j/a_i) \log p(b_j/a_i) = -\sum_i 0 \log 0 = 0,$$

аналогично

$$H(A/b_j) = -\sum_i p(a_i/b_j) \log p(a_i/b_j) = 0,$$

а, следовательно, и общая условная энтропия превращается в нуль и выражение для $H(A, B)$ приобретает вид (13.56).

Выводы: *1. Энтропия объединенной системы A, B равна безусловной энтропии одной из них плюс условная энтропия другой относительно первой.*

2. Матрица «объединения», описывающая взаимодействие систем или ансамблей сообщений при помощи вероятностей совместных событий, обладает свойством информационной полноты.

3. Взаимная энтропия ансамблей произвольных выборочных пространств обладает свойством взаимной симметрии.

4. В случае статистической независимости элементов ансамблей взаимная энтропия любого количества ансамблей равна сумме их безусловных энтропии.

5. При полной статистической зависимости ансамбля источника сообщений и ансамбля приемника сообщений их взаимная энтропия равна безусловной энтропии источника сообщений.

13.7. Энтропия непрерывного источника информации

Нами была рассмотрена мера неопределенности выбора для дискретного источника информации. На практике мы в основном встречаемся с источниками информации, множество возможных состояний которых составляет континуум. Такие источники называют *непрерывными* источниками информации. Во многих случаях они преобразуются в дискретные посредством использования устройств дискретизации и квантования. Вместе с тем существует немало и таких систем, в которых информация передается и преобразуется непосредственно в форме непрерывных сигналов. Примерами могут служить системы телефонной связи и телевидения.

Оценка неопределенности выбора для непрерывного источника информации имеет определенную специфику. Во-первых, значения, реализуемые источником, математически отображаются непрерывной случайной величиной. Во-вторых, вероятности значений этой случайной величины не могут использоваться для оценки неопределенности, поскольку в данном случае вероятность любого конкретного значения равна нулю.

Естественно, однако, связывать неопределенность выбора значения непрерывной случайной величины с плотностью распределения вероятностей этих значений. Учитывая, что для совокупности значений, относящихся к любому сколь угодно малому интервалу непрерывной случайной величины, вероятность конечна, попытаемся найти формулу для энтропии непрерывного источника информации, используя операции квантования и последующего предельного перехода при уменьшении кванта до нуля.

С этой целью разобьем диапазон изменения непрерывной случайной величины *U,* характеризующейся плотностью распределения вероятностей *p(u),* на конечное число *n* малых интервалов шириной Δu (рис. 13.16).

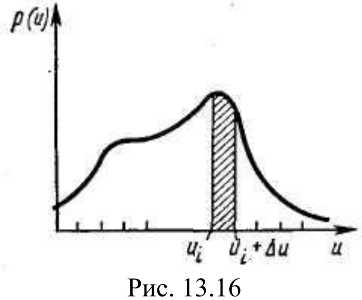


Рис. 13.16

При реализации любого значения *u*, принадлежащего интервалу $(u_i;\ u_i+\Delta u)$, будем считать, что реализовалось значение u_i дискретной случайной величины *U*. Поскольку Δu мало, вероятность $p(u_i \le u \le u_i+\Delta u)$ реализации значения *u* из интервала $u_i,\ u_i+\Delta u$:

$$p(u_i \leqslant u \leqslant u_i+\Delta u) = \int\limits_{u_i}^{u_i+\Delta u} p(u)du \approx p(u_i)\Delta u\,.$$

Тогда энтропия дискретной случайной величины *U* может быть записана в виде

$$H(\tilde{U}) = -\sum_{i=1}^{n} p(u_i)\Delta u \log\left[p(u_i)\Delta u\right],$$

или

$$H(\tilde{U}) = -\sum_{i=1}^{n} p(u_i)\Delta u \log p(u_i) - \sum_{i=1}^{n} p(u_i)\Delta u \log \Delta u\,.$$

Так как

$$\sum_{i=1}^{n} p(u_i)\Delta u = 1\,,$$

то

$$H(\tilde{U}) = -\sum_{i=1}^{n} p(u_i)\Delta u \log p(u_i) - \log \Delta u\,.$$

По мере уменьшения Δu $p(u_i \le u \le u_i+\Delta u)$ все больше приближается к вероятности $p(u_i)$, равной нулю, а свойства дискретной величины $\tilde{U}$ — к свойствам непрерывной случайной величины *U*.

Переходя к пределу при $\Delta u \to 0$, получаем следующее выражение для энтропии *H*(*U*) непрерывного источника:

$$\lim_{\Delta u\to 0} H(\tilde{U}) = \lim_{\Delta u\to 0}\left\{-\sum_{i=1}^{n} p(u_i)\Delta u \log p(u_i)\right\} - \lim_{\Delta u\to 0} \log \Delta u$$

или

$$H(U) = -\int\limits_{-\infty}^{\infty} p(u)\log p(u)du - \lim_{\Delta u\to 0} \log \Delta u\,. \qquad (13.57)$$

Эта величина при $\Delta u \to 0$ стремится к бесконечности, что полностью соответствует интуитивному представлению о том, что неопределенность выбора из бесконечно большого числа возможных состояний (значений) бесконечно велика.

Первый член в правой части соотношения (13.57) имеет конечное значение, которое зависит только от закона распределения непрерывной случайной величины *U* и не зависит от шага квантования

Δu. Он имеет точно такую же структуру, как энтропия дискретного источника.

Второй член того же соотношения, наоборот, зависит лишь от шага квантования случайной величины *U*. Именно в нем кроется причина того, что величина *H(U)* обращается в бесконечность.

К использованию и трактовке соотношения (13.57) для получения конечной характеристики информационных свойств непрерывного источника известны два подхода.

Один подход состоит в том, что в качестве меры неопределенности непрерывного источника принимают первый член соотношения (13.57):

$$h(U) = -\int_{-\infty}^{\infty} p(u)\log p(u)du\,. \tag{13.58}$$

Поскольку для определения этой величины используется только функция плотности вероятности, т. е. дифференциальный закон распределения, она получила название относительной дифференциальной энтропии или просто *дифференциальной энтропии* непрерывного источника информации (непрерывного распределения случайной величины *U).*

Ее можно трактовать как среднюю неопределенность выбора случайной величины *U* с произвольным законом распределения по сравнению со средней неопределенностью выбора случайной величины *U'*, изменяющейся в диапазоне, равном единице, и имеющей равномерное распределение.

Действительно, запишем соотношение (13.57) для случайной величины *U'*, равномерно распределенной в интервале δ:

$$H(U') = -\int_{-\infty}^{\infty} \frac{1}{\delta}\log\frac{1}{\delta}du' - \lim_{\Delta u'\to 0}\log\Delta u'\,.$$

При δ =1

$$H(U') = -\lim_{\Delta u'\to 0}\log\Delta u'\,,$$

откуда при $\Delta u = \Delta u'$

$$H(U) - H(U') = -\int_{-\infty}^{\infty} p(u)\log p(u)du = h(U)\,.$$

Аналогично, используя операции квантования и предельного перехода, найдем выражение для условной энтропии непрерывного источника информации:

$$H_V(U) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(u,v)\log\left[\frac{p(u,v)}{p(v)}\right] dudv - \lim_{\Delta u\to 0}\log\Delta u\,. \tag{13.59}$$

Отметим, что второй член в первой части выражения (13.59) идентичен соответствующему члену в соотношении (13.57). Обозначим первый член правой части выражения (13.59) через $h_V(U)$:

$$h_V(U) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(uv)\log\left[\frac{p(uv)}{p(v)}\right] dudv\,. \tag{13.60}$$

Эта величина конечна и называется относительной дифференциальной условной энтропией или просто *дифференциальной условной энтропией* непрерывного источника. Она характеризует неопределенность выбора непрерывной случайной величины U при условии, что известны результаты реализации значений другой статистически связанной с ней непрерывной случайной величины V, и по сравнению со средней неопределенностью выбора случайной величины U', изменяющейся в диапазоне, равном единице, и имеющей равномерное распределение вероятностей.

При втором подходе к использованию соотношения (13.57) для количественного определения информационных свойств непрерывного источника информации предлагается принять во внимание практическую невозможность обеспечения бесконечно высокой точности различения определенных значений непрерывной величины U. Поэтому все бесконечное число значений U в пределах заданной точности измерений следует рассматривать как одно значение.

Из средней неопределенности выбора источником u некоторого значения в этом случае необходимо вычесть среднюю неопределенность того же источника, полученную при условии, что мы знаем результаты определения U с некоторой определенной точностью ε. Тогда информационные свойства непрерывного источника будут оцениваться разностью безусловной и условной энтропии, определяемых соотношениями (13.57) и (13.59) соответственно. Такая разность, как будет показано дальше, является мерой снятой неопределенности, называемой *количеством информации.*

Таким образом, при втором подходе безусловная и условная энтропии непрерывного источника рассматриваются лишь как некоторые вспомогательные величины, с помощью которых можно определить количество информации. Соотношение между понятиями энтропии и количества информации для непрерывного источника информации подобно соотношению между потенциалом, определенным с привлечением понятия бесконечности, и напряжением, определенным как разность потенциалов.

Поскольку вторые члены в правых частях соотношений (13.57) и (13.59) одинаковы, разность безусловной и условной энтропии непрерывного источника информации равна разности дифференциальных безусловной и условной энтропии того же источника, причем относительность их уже несущественна, так как разность не зависит от стандарта, с которым они сравнивались.

Рассмотрим некоторые свойства дифференциальной энтропии.

1. Дифференциальная энтропия в отличие от энтропии дискретного источника является относительной мерой неопределенности. Ее значение зависит от масштаба случайной величины *U*, а следовательно, и от выбора единицы ее измерения.

Изменим масштаб случайной величины *U*, например, в *k* раз, оставив неизменным масштаб равномерно распределенной в единичном интервале случайной величины *U'*, принятой за эталон. Если $u_k = ku$, то $p(u_k) = p(u)/k$.

Тогда

$$h(u) = -\int_{-\infty}^{\infty} p(u_k)\log p(u_k)du_k =$$

$$= -\int_{-\infty}^{\infty} \frac{p(u)}{k}\log\left[\frac{p(u)}{k}\right]kdu = -\int_{-\infty}^{\infty} p(u)\times$$

$$\times\log p(u)du + \log k\int_{-\infty}^{\infty} p(u)du = h(u) + \log k. \tag{13.61}$$

Если одновременно изменить масштаб величины *U'*, то относительная неопределенность также изменится, так как значение эталона будет уже иным.

Из относительности дифференциальной энтропии следует, что энтропия может принимать положительные, отрицательные и нулевые значения.

2. Дифференциальная энтропия не зависит от конкретных значений случайной величины *U* и, в частности, от изменения всех ее значений на постоянное. Действительно, масштаб *U* при этом не меняется и справедливо равенство

$$h(U+\Theta)=-\int_{-\infty}^{\infty}p(u+\Theta)\log p(u+\Theta)d(u+\Theta)=$$

$$=-\int_{-\infty}^{\infty}p(u)\log p(u)du=h(U). \tag{13.62}$$

3. Какие же непрерывные распределения обладают максимальной дифференциальной энтропией?

а. Если единственным ограничением для случайной величины *U* является область ее возможных значений [α, β], то максимальной дифференциальной энтропией обладает равномерное распределение вероятностей в этой области.

При доказательстве решается задача определения плотности распределения *p(u),* обеспечивающей максимальное значение функционала

$$h(U)=-\int_{\alpha}^{\beta}p(u)\log p(u)du$$

при ограничении

$$\int_{\alpha}^{\beta}p(u)du=1. \tag{13.63}$$

Используя, например, метод неопределенных множителей Лагранжа, получим

$$p(u)=1/(\beta-\alpha),\ \ \alpha\leqslant u\leqslant\beta. \tag{13.64}$$

Нетрудно убедиться в том, что найденная функция *p(u)* обеспечивает максимум функционала *h(U)*, причем

$$h(U)=\log(\beta-\alpha). \tag{13.65}$$

б. Если ограничения на область значений непрерывной случайной величины *U* отсутствуют, но известно, что дисперсия ее ограничена, то максимальной дифференциальной энтропией обладает нормальное распределение величины *U*.

При доказательстве решается задача определения функции *p*(*u*), обеспечивающей максимальное значение функционала

$$h(U)=-\int_{-\infty}^{\infty}p(u)\log p(u)du \tag{13.66}$$

при ограничениях

$$\int_{-\infty}^{\infty}p(u)du=1\ \text{ и }\ \int_{-\infty}^{\infty}u^2p(u)du=\sigma^2,$$

где σ — среднеквадратическое отклонение от математического ожидания ($\tilde{U}$ = 0 (σ — заданное ограничение).

Искомую плотность распределения *p(u)* находят, например, методом неопределенных множителей Лагранжа.

Она оказывается гауссовской:

$$p(u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2\sigma^2}. \tag{13.67}$$

Вычислив функционал (13.66) для этого распределения, получим значение максимальной дифференциальной энтропии $h_{\max}(U)$. В двоичных единицах неопределенности

$$h_{\max}(U) = \log_2 \sigma\sqrt{2\pi e}. \tag{13.68}$$

Поскольку в информационных системах сигнал, описываемый случайной величиной *U,* часто представляет собой электрическое напряжение (или ток), дисперсия *U* пропорциональна средней мощности сигнала. Тогда в соответствии с (13.67) можно утверждать, что при заданной мощности наибольшей средней неопределенностью выбора будет обладать источник, генерирующий сигналы, амплитуды которых распределены по нормальному закону.

4. Соотношения для дифференциальной энтропии объединения статистически зависимых непрерывных источников аналогичны соответствующим формулам для дискретных источников:

$$h(UV) = h(U) + h_U(V) = h(V) + h_V(U), \tag{13.69}$$

где

$$h(UV) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(uv) \log p(uv) du dv.$$

Справедливость соотношения (13.69) легко проверить подстановкой выражения (13.58), заданного для *h(V),* и выражения (13.60) —для $h_V(U)$.

Так как

$$h_V(U) \leqslant h(U) \text{ и } h_U(V) \leqslant h(V),$$

то

$$h(UV) \leq h(U) + h(V),$$

причем равенство имеет место только в случае отсутствия статистической связи между *U* и *V.*

Пример. Определить, насколько мы выиграем в мощности, используя для организации мешающего воздействия, характеризующегося энтропией, источник шума с гауссовской плотностью распределения по сравнению с источником, имеющим в интервале [α, β] равномерную плотность распределения.

В соответствии с (13.68) дифференциальная энтропия гауссовского распределения

$$h_{\max}(u) = \log \sigma_г \sqrt{2\pi e},$$

где $\sigma_г^2$ — дисперсия, характеризующая мощность, выделяемую на резисторе с сопротивлением в 1 Ом.

Для равномерного распределения энтропия определена соотношением (13.65):

$$h_{\max}(U) = \log (\beta - \alpha).$$

Вычислим дисперсию $\sigma_р^2$ равномерного на интервале [α, β] распределения:

$$\sigma_р^2 = \int_\alpha^\beta \left(u - \frac{\alpha+\beta}{2} \right)^2 \frac{1}{\beta-\alpha} du = \frac{(\beta-\alpha)^2}{12}. \qquad (13.70)$$

Из условия обеспечения равенства энтропии следует

$$\log \sigma_г \sqrt{2\pi e} = \log(\beta - \alpha), \qquad (13.71)$$

$$\sigma_г \sqrt{2\pi e} = \beta - \alpha. \qquad (13.72)$$

Возведя (13.72) в квадрат и подставив в (13.70), получим

$$\sigma_р^2 = \frac{\pi e}{6} \sigma_г^2 = 1{,}42\, \sigma_г^2.$$

Следовательно, искомый выигрыш составляет 42 %.

13.8. Количество информации как мера снятой неопределенности

Передача информации инициируется либо самим источником информации, либо осуществляется по запросу. Она диктуется желанием устранить неопределенность относительно последовательности состояний, реализуемых некоторым источником информации. Обычно запрос обусловлен отсутствием возможности наблюдать состояния источника непосредственно. Поэтому абонент обращается к информационной системе, которая извлекает интересующую его информацию из источника посредством некоторого первичного преобразователя и направляет ее по каналу связи абоненту.

Информация проявляется всегда в форме сигналов. Сигналы *z*, поступающие с выхода первичного преобразователя источника информации на вход канала связи, принято называть сообщениями в отличие от сигнала *u*, формирующихся на входе линии связи. В зависимости от формы создаваемых сообщений различают источники дискретных и непрерывных сообщений.

Отдельные первичные сигналы с выхода источника дискретных сообщений называют *элементами сообщения.* Каждому элементу сообщения соответствует определенное состояние источника информации. В случае параллельной реализации источником информации множества состояний, как это имеет место, например, в документах с печатным текстом, первичный преобразователь, в частности, обеспечивает их последовательное отображение элементами сообщения. Таким преобразователем может быть как автоматическое читающее устройство, так и человек.

Основное понятие теории информации — *количество информации* — рассматривается в данном пункте применительно к передаче отдельных статистически несвязанных элементов сообщения. Дискретный источник сообщений при этом полностью характеризуется ансамблем

$$Z = \begin{pmatrix} z_1 & \dots & z_i & \dots & z_N \\ p(z_1) & \dots & p(z_2) & \dots & p(z_N) \end{pmatrix},$$

а непрерывный — одномерной плотностью распределения случайной величины — z — $p(z)$. Особенности определения количества информации при передаче сообщений изложены дальше.

Передача информации от дискретного источника. Выясним, насколько будет изменяться неопределенность относительно состояния источника сообщения при получении адресатом элемента сообщения с выхода канала связи. Алфавиты передаваемых и принимаемых элементов сообщения считаем идентичными.

Вследствие воздействия помех полученный элемент сообщения в общем случае отличается от переданного. Подчеркнем это различие. Обозначив принимаемые элементы сообщения другими буквами: $w_1, \dots.. \; w_j, \; \dots, \; w_N$.

Априорная неопределенность (неопределенность до получения элемента сообщения) относительно состояния источника не является полной. Предполагается, что адресату известен алфавит элементов сообщения, а из прошлого опыта он знает вероятности их появления. Считая, что состояния источника реализуются независимо, априорная частная неопределенность появления элемента сообщения z_i

$$H(z_i) = -\log p(z_i),$$

где $p(z_i)$ — априорная вероятность появления элемента сообщения z_i.

Педполагаются также известными некоторые сведения относительно помехи в канале связи. Обычно считают, что между элементами сообщения и помехой статистические связи отсутствуют, искажения отдельных элементов сообщения являются событиями независимыми и адресату известна совокупность условных вероятностей $p(z_i /w_j)$ $(1 \leq i \leq N,\ 1 \leq j \leq N)$ того, что вместо элемента сообщения z_i будет принят элемент сообщения w_j.

При получении конкретного элемента сообщения w_j адресату становится известным значение условной вероятности $p(z_i/w_j)$, называемой апостериорной (послеопытной) вероятностью реализации источником элемента сообщения z_i. Это позволяет найти апостериорную частную неопределенность, остающуюся у адресата относительно выдачи источников элемента сообщения z_i после получения конкретного элемента сообщения w_j.

$$H(z_i/ w_j) = - \log p(z_i /w_j). \qquad (13.73)$$

Поскольку получение информации мы связываем с уменьшением неопределенности, естественно определить частное количество информации $I(z_i)$, получаемое при приеме элемента сообщения w_j относительно некоторого реализованного источником элемента сообщения z_i как разность частных неопределенностей, имевшихся у адресата до и после получения элемента сообщения (априорной и апостериорной):

$$I(z_i) = H(z_i) - H(z_i/w_j) = \log \frac{p(z_i/w_j)}{p(z_i)} . \qquad (13.74)$$

Анализ формулы (13.74) позволяет сделать следующие заключения:

1) частное количество информации растет с уменьшением априорной и увеличением апостериорной вероятностей реализации элемента сообщения источником, что находится в полном соответствии с нашими интуитивными представлениями;

2) частное количество информации об элементе сообщения z_i может быть не только положительным, но и отрицательным, а также нулем, что зависит от соотношения априорной $p(z_i)$ и апостериорной $p(z_i/ w_j)$ вероятностей. Если вероятность того, что источником был реализован элемент сообщения z_i увеличилась после приема элемента сообщения w_j, т. е. $p(z_i/w_j)>p(z_i)$, то полученное частное количество информации положительно. Если эта вероятность не изменилась, т. е. $p(z_i /w_j) = p(z_i)$, то имевшая место неопределенность тоже не изменилась и частное количество информации равно нулю.

Наконец, случай $p(z_i/w_j)<p(z_i)$ соответствует увеличению неопределенности относительно реализации z_i после получения элемента сообщения w_j, и, следовательно, частное количество информации отрицательно;

3) в случае отсутствия помехи апостериорная вероятность $p(z_i/w_j)=1$. При этом частное количество информации численно

совпадает с частной априорной неопределенностью реализации данного элемента сообщения z_i:

$$I(z_i) = H(z_i) = -\log p(z_i).$$

Это максимальное частное количество информации, которое можно получить об элементе сообщения z_i;

4) частное количество информации относительно реализации источником элемента сообщения z_i, содержащееся в принятом элементе сообщения w_j, равно частному количеству информации относительно w_j, содержащемуся в элементе сообщения z_i

$$I(z_i w_j) = \log\frac{p(z_i/w_j)}{p(z_i)} = \log\frac{p(z_i w_j)}{p(z_i)p(w_j)} = I(w_j z_i). \tag{13.75}$$

Хотя имеют место случаи, где важно оценить частное количество информации $I(z_i w_j)$, для задач анализа и оптимизации функционирования информационных систем более рациональны усредненные характеристики, отражающие статистические свойства источника информации и канала связи.

Найдем среднее количество информации, содержащееся в любом принятом элементе сообщения относительно переданного (реализованного) источником. До получения конкретного элемента сообщения средняя неопределенность, имеющаяся у адресата, относительно реализации источником любого элемента сообщения равна энтропии источника. Ее называют *априорной энтропией* источника.

Средняя неопределенность относительно любого состояния источника, остающаяся у адресата после получения конкретного элемента сообщения w_j, характеризуется частной условной энтропией $H_{w_j}(z)$:

$$H_{w_j}(Z) = -\sum_{i=1}^{N} p(z_i/w_j)\log p(z_i/w_j). \tag{13.76}$$

Это случайная величина, зависящая от того, какой конкретно элемент сообщения принят.

Средняя неопределенность по всему ансамблю принимаемых элементов сообщений равна условной энтропии источника $H_W(Z)$:

$$H_W(Z) = \sum_{j=1}^{N} p(w_j) H_{w_j}(Z),$$

или

$$H_W(Z) = -\sum_{j=1}^{N}\sum_{i=1}^{N} p(w_j)p(z_i/w_j)\log p(z_i/w_j). \tag{13.77}$$

Эту условную энтропию называют *апостериорной энтропией* источника информации.

Таким образом, при наличии помех среднее количество информации, содержащееся в *каждом* принятом элементе сообщения, относительно *любого* переданного равно разности априорной и апостериорной энтропии источника:

$$I(Z) = H(Z)-H_W(Z). \qquad (13.78)$$

Представив априорную и апостериорную энтропии соответственно выражениями (13.6) и (13.77) и проведя несложные преобразования, получим формулу для количества информации непосредственно через вероятности:

$$I(ZW) = -\sum_{i=1}^{N} p(z_i)\log p(z_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} p(z_i w_j)\log p(z_i/w_j) =$$

$$= -\sum_{i=1}^{N}\sum_{j=1}^{N} p(z_i w_j)\log p(z_i) - \sum_{i=1}^{N}\sum_{j=1}^{N} p(z_i w_j)\log p(z_i/w_j) =$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} p(z_i w_j)\log\frac{p(z_i w_j)}{p(z_i)pw_j}. \qquad (13.79)$$

Если частный характер количества информации специально не оговаривается, мы всегда имеем дело с количеством информации, приходящимся *в среднем на один элемент сообщения.* Поэтому указание об усреднении опускается.

Передача информации от непрерывного источника. Количество информации, получаемой от непрерывного источника по каналу с помехами, определяется так же, как в случае, рассмотренном выше, но с использованием понятия дифференциальной энтропии.

Для источника, имеющего непрерывное множество состояний, среднее количество информации, содержащееся в каждом принятом значении случайной величины *W* относительно переданного значения случайной величины Z, можно получить как разность априорной и апостериорной дифференциальных энтропии:

$$I(ZW) = h(Z)-h_W(Z). \qquad (13.80)$$

Соотношение (13.80) несложно выразить в виде, подобном (13.79):

$$I(ZW) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(zw)\log\frac{p(zw)}{p(z)p(w)}dz\,dw\,. \qquad (13.81)$$

Относительность дифференциальных энтропий в этом случае не принимается во внимание, поскольку количество информации не зависит от выбранного стандарта сравнения.

Основные свойства количества информации. 1. Несмотря на то что частное количество информации может быть величиной отрицательной, количество информации неотрицательно.

Действительно, согласно выражению

$$H_W(Z) \leq H(Z). \qquad (13.82)$$

Тогда

$$I(ZW) = H(Z) - H_W(Z) \geq 0 .$$

2. При отсутствии статистической связи между случайными величинами Z и W

$$H_W(Z) = H(Z), \qquad (13.83)$$

следовательно, в этом случае

$$I(ZW) = 0$$

(принятые элементы сообщения не несут никакой информации относительно переданных).

3. Количество информации в W относительно Z равно количеству информации в Z относительно W.

Для доказательства этого утверждения воспользуемся выражением

$$H(ZW) = H(Z) + H_Z(W).$$

Аналогично можно записать

$$H(WZ) = H(W) + H_W(Z).$$

Так как $H(ZW) = H(WZ)$, то

$$H(Z) + H_Z(W) = H(W) + H_W(Z),$$

откуда

$$H(Z) - H_W(Z) = H(W) — H_z(W), I(ZW) = I(WZ) \qquad (13.84)$$

4. При взаимно однозначном соответствии между множествами передаваемых и принимаемых элементов сообщений, что имеет место в отсутствии помехи, апостериорная энтропия равна нулю и количество информации численно совпадает с энтропией источника:

$$I(ZW) = H(Z). \qquad (13.85)$$

Это максимальное количество информации о состоянии дискретного источника. Для непрерывного источника оно равно бесконечности.

Пример. Выстрел из орудия не поражает цель с вероятностью p. Через какое число выстрелов следует поинтересоваться у разведчика-корректировщика, уничтожена ли цель, чтобы в результате ответа получить максимальное количество информации?

Ансамбль интересующих нас событий включает: z_1 — цель поражена; z_2 — цель не поражена. Вероятность того, что цель не поражена после k выстрелов, равна p^k. Вероятность противоположного события $(1— p^k)$. Поскольку после ответа корректировщика неопределенность устраняется полностью, количество информации равно энтропии, а она максимальна при равновероятности событий. Следовательно,

$$p^k = 1 — p^k$$

откуда

$$k = -1/\log_2 p.$$

Ранее было показано, что неопределенность реализации непрерывным источником информации состояния в конкретный момент времени (отсчета) равна бесконечности. Тем более равна бесконечности неопределенность реализации непрерывным источником конкретного сигнала длительности *T*.

Однако такой результат получен в предположении возможности фиксировать любые сколь угодно малые различия между реализациями. На практике такая возможность отсутствует. Это объясняется тем, что воспринимающие информацию датчики, включая человека, обладают ограниченной чувствительностью и конечной разрешающей способностью, а также тем, что процесс восприятия сопровождается помехами.

Если учесть, что нас интересует приближенное восприятие реализации, то количество информации, приходящееся на отсчет или на единицу времени, можно вычислить.

Ограничимся рассмотрением простейшего случая, когда отдельные состояния источника информации представляют собой независимые реализации случайной величины *U*. (Эпсилон-энтропия случайного процесса рассмотрена дальше.)

Ансамбль реализаций случайной величины *U* описывается плотностью распределения вероятностей *p(u)*. О значениях случайной величины *U* можно судить по значениям другой случайной величины Z, если мера их различия не превышает заданной верности воспроизведения. В этом случае говорят, что Z воспроизводит *U*.

Для количественной оценки степени сходства сигналов целесообразно ввести какую-либо функцию ρ(*z, u)*, имеющую природу «расстояния». Тогда удобным критерием верности *V(Z, U)* является среднее значение функции ρ(z*, u)*, взятое по всему множеству значений *z* и *u:*

$$V(ZU) = \iint p(zu)\rho(zu)\,dzdu\,, \qquad (13.86)$$

где ρ(*z, u)* — плотность совместного распределения вероятностей величин *Z* и *U*.

Наиболее широко используется среднеквадратический критерий, при котором ρ(z*, u)* представляет собой квадрат обычного евклидова расстояния между точками в соответствующем пространстве

Требование к верности в данном случае задается с использованием критерия $\tilde{V}$ *(ZU):*

$$\tilde{V}(ZU) = \iint p(u) p_u(z)(z-u)^2 du dz \leqslant \varepsilon^2, \quad (13.87)$$

где $p_u(z)$ — условная плотность распределения — функция правдоподобия того, что конкретный сигнал u будет воспроизведен как сигнал z; ε — заданное значение верности.

Так как плотность $p(u)$ определена, то для выполнения условия (13.87) варьировать можно только условной плотностью распределения $p_u(z)$.

Если случайная величина Z воспроизводит случайную величину U с некоторой верностью ε, то количество информации, содержащееся в воспроизводящей величине Z относительно U, конечно и в соответствии с (13.81) может быть записано в форме

$$I(ZU) = \iint p(u) p_u(z) \log \frac{p_u(z)}{p(z)}. \quad (13.88)$$

где $p(z) = \int p(u)p_u(z)du$ — плотность воспроизводящей величины Z.

Желательно обеспечить заданную верность воспроизведения при *минимальном* количестве получаемой информации. Поэтому среди множества функций $p_u(z)$, удовлетворяющих условию (13.87), целесообразно выбрать такую, которая обеспечивает наименьшее $I(ZU)$.

Минимальное количество информации в одной случайной величине Z относительно другой U, при котором удовлетворяется заданное требование к верности воспроизведения величины U, называется ε-энтропией величины U и обозначается $H_\varepsilon(U)$:

$$H_\varepsilon(U) = \min_{\{p_u(z)\}} I(ZU) \quad (13.89)$$

при

$$V(ZU) \leqslant \varepsilon^2. \quad (13.90)$$

Используя безусловную $h(U)$ и условную $h_z(U)$ дифференциальные энтропии величины U, выражение (13.90) можно представить в виде

$$H_\varepsilon(U) = h(U) - \max_{\{p_z(u)\}} h_Z(U), \quad (13.91)$$

где $p_z(u)$ — условная плотность вероятности того, что в тех случаях, когда был принят сигнал z, передавался сигнал u.

Пример. Найти $H_\varepsilon(U)$ источника информации, ансамбль состояний которого описывается нормально распределенной случайной величиной U с дисперсией σ^2 при верности воспроизведения

$$V'(ZU) \le \varepsilon^2.$$

Будем считать, что заданная верность воспроизведения обусловлена действием аддитивной статистически не связанной с сигналом помехой

Ξ, причем $M[\Xi]=0$ и $M[\Xi^2]=\varepsilon^2$. Передаваемый сигнал u рассматриваем как сумму воспроизводящего сигнала z и помехи $u=z+\xi$

Так как в данном случае $h_z(U)$ в выражении (13.91) полностью определяется помехой $[h_z(U)=h_z(Z+\Xi)=h(\Xi)]$, то

$$H_\varepsilon(U)=h(U)-\max_{p(\xi)} h(\Xi),$$

где $h(\Xi)$— дифференциальная энтропия помехи; $p(\xi)$ — плотность распределения помехи Ξ.

Ранее [см. (13.67)] нами установлено, что при ограничении на дисперсию случайной величины максимальной дифференциальной энтропией обладает нормальное распределение. Поэтому в соответствии с (13.68) получаем

$$\max_{\{p(\xi)\}} h(\Xi)=\log \varepsilon\sqrt{2\pi e},$$

$$H(U)=\log \sigma\sqrt{2\pi e},$$

откуда

$$H_\varepsilon(U)=\log \sigma\sqrt{2\pi e}-\log \varepsilon\sqrt{2\pi e}=\frac{1}{2}\log\frac{\sigma^2}{\varepsilon^2}. \tag{13.92}$$

Так как σ^2 определяет среднюю мощность P_u сигнала, а ε^2 — среднюю мощность p_ξ помехи Ξ, то выражение (13.92) характеризует зависимость эпсилон-энтропии от величины P_u/P_ξ, называемой *отношением сигнал/помеха.*

При заданном отношении сигнал/помеха значение $H_\varepsilon(U)$ для нормально распределенной случайной величины является максимально возможным.

Для произвольно распределенной случайной величины U при том же критерии верности и малых ε [когда $H_\varepsilon(U)$ велико] справедливо приближенное равенство

$$H_\varepsilon(U)\approx h(U)-\log \varepsilon\sqrt{2\pi e}. \tag{13.93}$$

13.9. Энтропия как мера неупорядоченности статистических форм движения

Решаемые с помощью теории информации задачи техники связи в значительной мере усложняются тем, что далеко не вся переданная информация доходит до получателя сообщений, так как значительная часть ее теряется в канале связи из-за воздействия помех. Главная задача, которую решает теория информации в ее, так сказать, «классическом» виде, заключается именно в том, чтобы с помощью специально разработанных кодов и оптимально найденных параметров каналов связи свести к минимуму значения обусловленных шумами информационных потерь.

С точки зрения интересующего нас аспекта объективной роли информации в эволюционных процессах важно обратить внимание на то обстоятельство, что в «классической» теории информации одним из звеньев канала передачи и приема информации является ее получатель, субъективные свойства которого (априорная и апостериорная осведомленность) играют при исчислениях информации принципиальную роль. Исходя из этого, как мы уже говорили, предлагалось определить количественную меру информации как «меру неведения» или «меру неожиданности», подчеркивая этим тот факт, что чем неожиданнее (малонероятнее) для получателя сообщение, тем больше информации оно несет. (В качестве иллюстрации обратного явления можно вспомнить пример сообщений, подобных фразе «Волга впадает в Каспийское море». Благодаря априорной осведомленности об этом факте, сообщение не является неожиданным для большинства получателей, и потому для них содержащаяся в нем информация равна нулю.)

Р. Хартли, который предпринял первую попытку исчисления информации, старался полностью исключить фактор «неожиданности». Как мы уже говорили ранеее, он предложил определять количество информации с помощью формулы: $I = \log N$, где N — число возможных вариантов сообщения (например, число букв алфавита).

Хартли понимал, что разные сообщения имеют различную вероятность, а потому неожиданность их появления для получателя различна. Однако он решил, что фактор «неожиданности» «относится к компетенции психологии и не должен интересовать инженера-связиста». К. Шеннон ввел в теорию «неожиданность», предложив учитывать различные вероятности с помощью функции

$$\sum_i p_i \log p_i.$$

Оказалось, что формула, предложенная Хартли,— это всего лишь частный случай более общей формулы Шеннона. В самом деле, если в формуле Шеннона принять

$$p_1 = p_2 = \ldots = p_i = \ldots p_N = \frac{1}{N},$$

мы получаем:

$$I = -\sum_{i=1}^{i=N} \frac{1}{N} \log \frac{1}{N} = -\log \frac{1}{N} = \log N.$$

Тот факт, что формула Шеннона в самом деле учитывает неожиданность сообщений, подтверждается не только общими рассуждениями, но и конкретным психологическим экспериментом, о котором будет рассказано ниже. Но означает ли это, что предложенная *К.* Шенноном мера количества информации в основе своей субъективна? Если бы это было так, понятие «информации» никогда не смогло бы выйти за рамки своего первоначального обиходного смысла и оставалось бы термином, обозначающим приобретение тем или иным субъектом каких-то сведений или знаний. Между тем, по мере развития науки и техники выявляется все большее количество информационных процессов, протекающих без участия каких бы то ни было наблюдателей и субъектов. Примером тому может служить обмен информацией между живыми (но отнюдь по мыслящими, а стало быть, и не осознающими информацию и ее «неожиданность») клетками, между звеньями автоматов и электронных машин. В науке возникла настоятельная потребность разграничить объективное и субъективное в содержании понятия информации, рассмотреть и те случаи, когда количество информации в самом деле определяется осведомленностью получателя, и те случаи, когда субъект из процесса полностью исключен.

Собственно говоря, с момента рождения теории информации этот вопрос все время стоит на повестке дня. Именно он заставил Р. Хартли исключить, из формулы «неожиданность»; именно он заставил К. Шеннона неоднократно подчеркивать, что предложенная им мера количества информации учитывает лишь вероятности сообщений, но игнорирует такие субъективные характеристики сообщений, как их ценность и смысл. Согласно теории Шеннона, сообщение о том, что родился мальчик или девочка, содержит такое же количество информации, как сообщение о том, кому из шахматных партнером выпал первый ход белыми (1 бит), хотя эти сообщения имеют совершенно разную субъективную ценность для отца родившегося ребенка и для «болельщика» шахмат.

Однако исключение ценности и смысла информации устранило субъективизм этого понятия не полностью, а лишь частично. Дело в том, что входящие в функцию $\sum_i p_i \log p_i.$ вероятности тоже в той или иной степени субъективны. Так же, как и информация, вероятность может характеризовать и объективные свойства стохастичтшх процессов, и степень недостатка исходных сведений, когда, например, в силу незнания статистических свойств исследуемой системы наблюдатель первоначально считает, что вероятности всех исходов равны (как это и имело место в рассмотренном выше примере письменных текстов, когда не известны вероятности отдельных буки). Для того чтобы из априорно заданных значений вероятностей полностью исключить обусловленную первоначальной неосведомленностью субъективность и получить значение вероятностей, характеризующее объективные статистические свойства исследуемых систем, требуется получить и обработать огромное количество данных. На вопросе об объективной и субъективной сущности понятия вероятности мы остановимся подробнее ниже. Задачей данного раздела является выявление объективного содержания понятия информации и его связи с понятием энтропии.

Развитие кибернетики со всей очевидностью показало, что способность сохранения априорных и апостериорных сведений (т. е. информации предварительной и вновь поступившей) о том или ином объективном явлении присуща не только обладающим сознанием субъектам, но и искусственно созданным системам памяти электронных машин. Основой создания искусственной памяти является в свою очередь способность тех или иных физических тел сохранять (запоминать) результаты воздействия электрических и магнитных полей.

Очевидно, что понятие «неожиданность», вполне приемлемое при рассмотрении получающего информацию субъекта, становится неприменимым, когда речь идет об объективном взаимодействии неодушевленных (а потому ничего не «ожидающих») физических тел. Однако и в этом случае процесс получения и накопления информации может быть описан с помощью той же функции $\sum_i p_i \log p_i$.

Чтобы убедиться в этом, достаточно представить себе систему, элементы которой двигались бы беспорядочно (т. е. имели равные вероятности всех направлений), а затем, в силу каких-то внешних причин, приобрели упорядоченность движения, характеризуемую неравномерным распределением значений p_i. Это значит, что до воздействия система обладала энтропией $H_1 = H_{\max}$, а после воздействия ее энтропия уменьшилась до величины H_2. Другими словами, благодаря воздействию извне система приобрела информацию $I_0 = H_1 — H_2$. Именно в этом и заключается сущность установленного Л. Бриллюэном негэнтропийного принципа информации, согласно которому, приобретая информацию H_0, система неизбежно уменьшает свою энтропию на величину ΔH, причем $|\Delta H| = I_0$.

В этой связи хотелось бы подчеркнуть, что понятие информации совершило в науке своеобразный и необычный путь. Сначала из обиходного термина, означавшего приобретение новых сведений и знаний, информация превратилась в строго научное понятие, имеющее свою количественную меру и пригодное для строгого и точного анализа сложных каналов связи и различных кибернетических систем. Затем понятие информации распространилось на организмы и клетки. Негэнтропийный принцип информации позволяет распространить понятие информации на область исследования кристаллов, молекул, атомов и прочих упорядоченных неорганических систем. Очевидно, настало время, когда, отказавшись от хронологической последовательности событий, можно попытаться поставить этот вопрос «с головы на ноги», начав информационные исследования не с мозга и электронной машины, а с таких простых систем, как газ, жидкость или кристалл. Для этого придется прежде всего рассмотреть с новых позиций сущность понятия физической энтропии.

Рассмотрение сложной картины явлений, связанных с изменением энтропии различных материальных систем, следует начать с установления связи энтропии системы с характером движения элементов системы. Как известно, состояние термодинамического равновесия физических тел (в частности, газа) характеризуется максимальным значением энтропии, означающим наибольшую неупорядоченность движения молекул. Данному состоянию соответствует наиболее равномерное распределение значений скоростей. Абстрагируясь от конкретных объектом, можно представить себе такую систему, в которой в каком-то интервале скоростей перемещение каждой молекулы в любом направлении и с любой скоростью имело бы одинаковую вероятность. В этом,случае согласно рассмотренным свойствам функции $\sum_i p_i \log p_i.$ определяемая формулой энтропия *H* достигает максимального значения.

Однако статистический анализ отвергает возможность реального существования подобных систем. Как известно, плотность вероятности, соответствующая наибольший неопределенности системы при заданной мощности процесса, подчиняется нормальному закону распределения. Указанное свойство находит свое отражение в кривых Максвелла, приведенных на рис. 13.17.

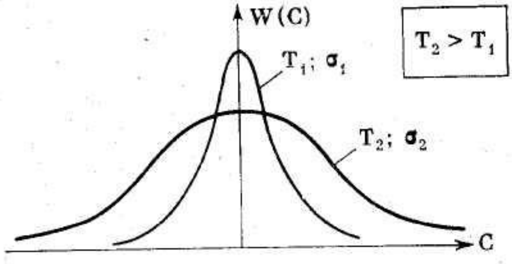


Рис. 13.17

Сопоставляя кривые распределения скоростей при различных температурах (T_1 и T_2 на рис. 13.17), нетрудно прийти к заключению, что увеличение энтропии идеального газа при увеличении его температуры от T_1 до T_2 обусловлено увеличением неупорядоченности движения молекул, т. е. более равномерным, приближающимся к равновероятному распределением скоростей *W (С)*.

Увеличение энтропии, происходящее при нагревании тела от температуры T_1 до температуры T_2, определяется классической термодинамикой как:

$$\Delta H = \frac{1}{K}\int_{T_1}^{T_2}\frac{Q}{T}\,dT, \tag{13.94}$$

где K — постоянная Больцмана; Q — количество тепла, необходимого для нагревания газа от температуры T_1 до температуры T_2. При этом классическая термодинамика не рассматривает изменений микроструктуры нагреваемого тела. Соотношение (13.94) констатирует лишь тот факт, что при получении физическим телом определенного количества тепла его энтропия растет.

Статистическая физика позволяет связать увеличение энтропии идеального газа с увеличением неупорядоченности движения молекул, выражающимся увеличением дисперсии кривой распределения вероятностей их скоростей (рис. 13.17). Согласно разработанной Дж. Максвеллом кинетической теории идеального газа в состоянии термодинамического равновесия распределение вероятностей скоростей подчиняется нормальному закону. При этом дисперсия вероятностного распределения связана с температурой газа соотношением:

$$\sigma = \sqrt{\frac{KT}{m}}, \tag{13.95}$$

где m — масса молекулы.

Изестно, что при нормальном законе распределения энтропия выражается как:

$$H = \ln \sigma \sqrt{2\pi} + \frac{1}{2}. \tag{13.96}$$

Используя соотношения (13.95) и (13.96), нетрудно определить приращение энтропии, происходящее при нагревании идеального газа от температуры T_1 до температуры T_2 (рис. 13.17):

$$\Delta H = \ln \frac{\sigma_2}{\sigma_1} = \frac{1}{2} \ln \frac{T_2}{T_1} \tag{13.97}$$

(Соотношение (13.97) не учитывает направления движения молекул. Учет распределения скоростей по трем осям коордишгг приводит к выражению:

$$\Delta H(x; y; z) = \frac{3}{2} \ln \frac{T_2}{T_1}.$$
)

Выражение (13.97) и представленные на рис. 13.17 кривые Максвелла показывают, что приращение энтропии идеального газа обусловлено увеличением дисперсии распределения вероятностей, т. е. более неупорядоченным, приближающимся к равновероятному распределением скоростей.

Кривые Максвелла отражают только распределение абсолютных значений скорости. Понятие энтропии является более «емким»: энтропия зависит и от абсолютных значений скоростей, и от направлений движения, и от пространственного распределения движущихся частиц. С целью статистического учета всех параметров движения множества частиц Дж. Гиббс предложил использовать многомерное фазовое пространство (Γ-пространство), с числом координатных осей, равным $2mN$, где N — число входящих в рассматриваемую систему частиц, а m — число степеней свободы каждой частицы. В результате такого анализа Гиббс получил следующее выражение энтропии:

$$H = -\frac{1}{\text{K}} \ln \int d\omega, \tag{13.98}$$

где $d\omega$ — элементарный объем фазового пространства, определяемый как:

$$d\omega = dv_q dv_p; \tag{13.99}$$

$$dv_q = dq_1 dq_2 \ldots dq_N; \tag{13.100}$$

$$dv_p = dp_1 dp_2 \ldots dp_N; \tag{13.101}$$

q — геометрические координаты частиц в $Г$-пространстве; p — динамические координаты частиц (импульсы) в $Г$-пространстве.

Приведенные соотношения показывают, что энтропия системы растет по мере увеличения неупорядоченности движения элементов, так как при этом увеличивается диапазон возможных значений координат и импульсов, а следовательно, и объем фазового пространства, определяемый соотношением (13.99) и связанный с величиной энтропии соотношением (13.98).

Таким образом, проделанный Гиббсом анализ еще раз подтвердил, что состоянию термодинамического равновесия, характеризуемому максимальным значением энтропии, соответствует наибольшая неупорядоченность движения частиц. Неупорядоченность эта выражается максвелловским распределением скоростей, равной вероятностью всех направлений движения и приблизительно равным количеством частиц, находящихся в одинаковых объемах занимаемого газом пространства в каждый момент. Для наглядности можно противопоставить этому случаю такое упорядоченное движение, при котором все частицы движутся в одном направлении с равными скоростями, соблюдая при этом, как на параде или в балете, определенный пространственный «узор». Этот образ введен нами чисто абстрактно, однако известно, что в природе существует немало подобных, в той или иной степени упорядоченных систем. Используя методы Больцмана или Гиббса, можно показать, что по мере увеличения упорядоченности движения элементов системы энтропия последней падает, стремясь в пределе к нулю (например, когда возможно только одно абсолютное значение и направление скорости, а вероятности остальных равны нулю).

Мы коснулись методов статистического анализа энтропийных процессов, разработанных в свое время Л. Больцманом и Д. Гиббсом, преследуя единственную цель: показать, что и многомерное фазовое пространство Гиббса и ранее разработанный Больцманом метод ячеек в пространстве скоростей, с помощью которого удалось впервые дать вероятностную интерпретацию энтропии, появились с целью статистического учета *пораметров движения микроэлементов* систем.

Однако принятые в статистической физике определения энтропии как «меры вероятности состояния» или «меры неопределенности состояния» физических тел не отражают взаимосвязи движения с энтропией. Рассмотренные соотношения позволяют сделать вывод о том, что физическая энтропия является *статистической мерой неупорядоченности движения микроэлементов физических тел.*

Определение энтропии как меры неупорядоченности движения остается в силе и в том случае, когда от непрерывного распределения вероятностей $W(X)$ переходят к рассмотрению характерных для информационных процессов дискретных значений вероятностей p_i. Именно благодаря этой общности понятия энтропии можно рассматривать процессы передачи и накопления информации в непосредственной связи с физическим состоянием участвующих в этих процессах материальных тел. Так же как и физическая энтропия, энтропия сообщений возрастает по мере увеличения равномерности распределения значений p_i. В предельных случаях вероятности всех значений параметров передаваемого сигнала (частот, уровней, яркостей и т. п.) будут равны друг другу. При этом звуковой сигнал становится шумом (например, так называемый Гауссов шум характеризуется равновероятным распределением всех частот в широким диапазоне, а распределение вероятностей амплитуд подчиняется нормальному (Гауссовому) закону), изображение представляет собой хаотичное чередование всех ступеней яркости в кадре и напоминает снежный буран. Энтропия таких сообщений достигает максимума, а количество информации равно нулю.

Текст с максимальной энтропией — это текст с равновероятным распределением всех букв алфавита, т. е. бессмысленное чередование букв. Вот пример подобного текста:

СУХРРОБЫЦ ЯЫХВЩИЮАЙЖТЛ-ТФВЛЭСТРОЕНВШТЦРИХГЪКУЧЖЮРЯНЧЬКЙХРЫС.

13.10. Информация как мера упорядоченности статистических форм движения

Если при составлении текста учтена реальная вероятность букв, в получаемых таким образом «фразах» уже наблюдается определенная упорядоченность движения (чередования) букв: ЕЫТ ЦИЯЬА ОКРВ ОДНТ ЬЧЕ МЛОЙК ЗЬЯ ЕНВ ТША.

При учете вероятностей четырехбуквенных сочетаний текст становится настолько упорядоченным, что по некоторым формальным признакам весьма напоминает осмысленный текст: ВЕСЕЛ ВРАТЬСЯ НЕ СУХОМ И НЕ-ПО И КОРКО.

В данном случае информация о статистических закономерностях письменных текстов является причиной упорядоченного движения (чередования) в создаваемой последовательности букв. В осмысленных фразах вероятности появления тех или иных букв еще более дифференцированы. В соответствии со свойствами функции

$$\sum_i p_i \log p_i$$

.

Это приводит к уменьшению энтропии текста,

$$|AH| = I.$$

Таким образом, *количеством информации,* содержащейся в том или ином сообщении, можно характеризовать *степень упорядоченности движения (чередования) кодовых знаков.* Упорядоченность будет тем больше, чем больше дифференцированы значения p_i. Благодаря этой дифференциации передаваемый радиосигнал сохраняет членораздельную речь или музыку, а на телеэкране возникает изображение, представляющее собой упорядоченное чередование различных ступеней яркости.

Следует еще раз подчеркнуть, что и в данном случае

функция $\sum_i p_i \log p_i$ опять-таки характеризует *движение.*

В общем случае в формуле Шеннона под символом p_i подразумевается вероятность каких-то событий, а события могут происходить лишь там, где существует движение. Передача любых сообщений осуществляется путем упорядоченного движения условных знаков, уровней, яркостей или частот. Для сохранения информации в ячейках памяти электронной машины необходимо поддерживать строго упорядоченное движение электронных потоков или магнитных полей.

С целью сохранения информации человек создал «неподвижные модели движения», к числу которых относятся кинопроекционные ленты, магнитные носители или печатный текст. Однако для получения информации от любой «неподвижной модели» необходимо вспомогательное движение: лента движется электромотором, движение букв, существовавшее в момент написания текста, восстанавливается глазами, движущимися вдоль строк. Информация, передаваемая с пульта управления автоматическому станку, содержит в себе «модель» (алгоритм) движения его органов. В то же время подача дополнительной энергии при отсутствии новых команд может лишь увеличить интенсивность движения, сохранив прежний порядок. Таким образом, *энергия является мерой интенсивности движения, а информация — мерой его упорядоченности.*

Как правило, все технические средства передачи информации предназначены для передачи «модели движения» на уровнях энергии значительно более низких, чем уровни энергии наблюдаемого движения или того движения, которым они должны управлять (в некоторых случаях анергия сигналов превышает энергию наблюдаемого движения ввиду того, что последняя слишком мала. Примером такого устройства может служить счетчик Гейгера — Мюллера.) Такое соотношении уровней энергии обеспечивает возможность управлять движением на расстоянии при условии минимальных потерь энергии в каналах связи.

В области биологических явлений получении информации является необходимым условием целесообразной направленности всех жизненных процессов. Различные внешние раздражители служат сигналом для включения тех или иных «моделей движения», сохраняемых центральной нервной системой живых организмов. Эти «модели» управляют рефлекторными движениями внешних и внутренних органов, обеспечивая тем самым рациональное взаимодействие со средой.

Используемая для оценки количества информации функция $\sum\limits_i p_i \log p_i$ позволяет оценивать степень упорядоченности движения, существующего в материальных системах, предназначенных для передачи или хранения информации. С таким же успехом с помощью данной функции можно оценивать упорядоченность движения в естественно протекающих физических процессах. В предыдущем пункте мы говорили о том, что тело, находящееся в состоянии термодинамического равновесия, характеризуется равномерным (точнее, Гауссовым) распределением вероятностей скоростей. При этом функция $\sum\limits_i p_i \log p_i$ достигает максимальной величины, а упорядоченность движения микроэлементов системы стремится к нулю. Но стоит лишь вывести систему из этого состояния, создав в различных частях ее разность температур или давлений, как в ней возникнут преобладающие направления движения в сторону более низких давлений и более низких температур. Степень упорядоченности движения, существующего в этой системе, в принципе может быть оценена с помощью той же функции $\sum\limits_i p_i \log p_i$. Для этого надо учесть, что в системе есть преобладающее направление, которому соответствует наибольшая вероятность p_i. Вероятность других направлений становится меньше (поскольку $\sum\limits_i p_i = 1$), а перераспределение вероятностей, как мы отмечали выше, приводит к уменьшению абсолютной величины функции $\sum\limits_i p_i \log p_i$. В ходе теплообмена (или выравнивания давлений) различие в величинах p_i становится меньше и энтропия $H = -\sum\limits_i p_i \log p_i$ будет расти.

Таким образом, использование одной и той же функции $\sum_i p_i \log p_i$ при анализе энтропии сообщений и энтропии физических тел отнюдь *не является чисто формальным математическим приемом,* так как и в том и в другом случае с помощью этой функции оценивается степень упорядоченности движения элементов тех или иных материальных систем.

Уменьшение абсолютной величины $\sum_i p_i \log p_i$, происходящее за счет перераспределения вероятностей p_i различных направлений движения, можно расценивать как возникновение *негэнтропии.* В соответствии с негэнтропийным принципом информации, установленным Л. Бриллюэном, количеству информации I, сохраняемой в системе памяти электронной машины, соответствует негэнтропия ΔH, возникающая в данной системе, причем $|\Delta H| = I$. Согласно Л. Бриллюэну, в равной степени справедливо и обратное утверждение: наличие негэнтропии в физическом теле, обладающем разностью температур или давлений, может расцениваться как сохранение информации в количестве $I = | \Delta H |$. Таким образом, для физических тел, так же как и для информационных процессов, функция $\sum_i p_i \log p_i$ позволяет оценивать степень упорядоченности движения, и лишь благодаря упорядоченности движения, присущей кристаллическим или магнитным ячейкам, мы можем для хранения информации прибегать к помощи физических тел. Увеличение степени упорядоченности движения является общей тенденцией эволюции материального мира. Без этой тенденции, проявляющейся даже в простых неорганических формах материи, невозможно было бы возникновение более сложных информационных систем.

13.11. Развитие — накопление информации

Наиболее ярким примером, иллюстрирующим связь эволюции материальных систем с увеличением степени упорядоченности движения, может служить живая природа. Весь процесс развития жизни и переход от простых ее форм к более сложным связаны с передачей наследственной информации и с закреплением рациональных наследственных признаков путем естественного отбора. В свете рассмотренных выше закономерностей есть основания утверждать, что информация, сохраняемая половой клеткой, содержит в себе «модель развития» вновь рожденного организма, подобно тому как информация, передаваемая с пульта управления, является «моделью движения» автоматического станка.

Диалектика издавна утверждает, что развитие — это одна из форм движения. Теперь к этому можно добавить, что информация — это единая мера упорядоченности движения, пригодная для оценки любых его форм, начиная от механических перемещений частиц в пространстве и кончая процессами развития самых сложных систем.

В самом деле, необходимым условием развития организма является непрерывное получение информации, с помощью которой осуществляется процесс авторегулирования, обеспечивающий устойчивость организма в различных условиях внешней среды. По мере развития форм живой материи усложнялась структура нервной системы и увеличивалось количество информации, сохраняемой организмами. Вместе с тем увеличивалась и упорядоченность движения, что приводило к более четкой дифференциации функций различных органов и усложнению общей структуры биологических систем.

Однако жизнь не могла возникнуть случайно: ее появление было подготовленным закономерным этапом в общем ходе развития материального мира. При этом общей тенденцией эволюции являлся процесс увеличения степени упорядоченности движения, присущий всей неорганической природе, начиная от процесса возникновения планетных систем и галактик и кончая процессами упорядочения движения, свойственными микроструктуре материальных систем. Отличие неупорядоченных физических систем (например, газа) от систем, обладающих упорядоченной структурой, заключается в том, что последние, находясь в состоянии устойчивого равновесия, сохраняют стабильные траектории движения элементов. Разумеется, указанную «стабильность» следует понимать в статистическом смысле, т. е. полагать, что определенные направления движения имеют в данном случае большую вероятность. Примером такой системы могут служить кристаллические структуры твердых веществ.

Перераспределение вероятностей направлений движения, возникающее при переходе от жидкой фазы к кристаллическому состоянию вещества, приводит к появлению негэнтропии, которая может быть оценена как уменьшение абсолютного значения величины $\sum_i p_i \log p_i$. При этом система приобретает способность передачи «моделей движения» окружающим молекулам жидкости. Такая «модель» передается от кристаллов, внесенных извне, в случае выращивания кристаллов, или от центров кристаллизации, возникающих в структуре кристаллизующихся веществ. Этот процесс передачи «моделей движения» является информационным процессом и переходным этапом к возникновению более сложных форм.

Науке известны различные стадии существования вирусов. В одной стадии вирус кристаллизуется, а в другой приобретает качественные признаки, присущие только живым организмам, в том числе и способность передачи «моделей движения», имеющих биологический код.

Сопоставив предлагаемую трактовку понятия информации с некоторыми выводами теории вероятностей, можно обосновать объективную необходимость увеличения упорядоченности движения в результате взаимодействий. Исследуя взаимодействие двух статистических множеств $(X$ и $Y)$, обладающих энтропией $H(X)$ и $H(Y)$, А. Я. Хинчин доказал, что результирующая энтропия $H(X, Y)$ удовлетворяет условию

$$H(X; Y) \leq H(X) + H(Y). \qquad (13.102)$$

Характеризуемое соотношением (13.102) взаимодействие, рассмотренное А. Я. Хинчиным в чисто математическом аспекте, приложимо и к случаям иааимодействий конкретных статистических материальных систем. Как показал А. Я. Хинчин, знак равенства в вырижеиии (13.102) означает независимость множеств X и Y. С физической точки зрения независимость означает отсутствие взаимодействий (например, система находящаяся в магнитном ноле, но не имеющая магнитных свойств). Во всех остальных случаях $H(X, Y)$ оказывается меньше, чем $H(X) + H(Y)$. Таким образом, соотношение (13.102) показывает, что *энтропия уменьшается в результате взаимодействий статистических систем.* Уменьшение энтропии эквивалентно росту негэнтропии, т. е. накоплению информации и увеличению степени упорядоченности системы, возникшей в результате взаимодействия двух исходных систем.

Многочисленные примеры из области термодинамики могут служить подтверждением тенденции к упорядочению движения и усложнению структуры в процессах взаимодействия физических систем. К числу таких процессов относятся: поляризация диэлектрика при взаимодействии с электрическим полем; переход теплового движения в упорядоченное движение кристаллической решетки при агрегатных превращениях веществ; образование веществ, обладающих сложной структурой молекул, в результате взаимодействия простых веществ (например:

$2Cu + S = Cu_2S + Q$, где Q — количество тепла).

Сопоставление этих разнородных процессов позволяет установить у них одно общее свойство: все они сопровождаются выделением тепла или аккумуляцией энергии внутренними связями. До сих пор данное свойство устанавливалось опытным путем. В настоящее время можно утверждать, что необходимость выделения тепла обусловлена накоплением информации, так как в соответствии со вторым началом термодинамики данное накопление должно быть компенсировано увеличением энтропии окружающей среды. В случае внутренней аккумуляции энергии прирост энтропии среды осуществляется за счет дополнительных внешних источников энергии.

К сожалению, еще невозможно производить количественную оценку негэтропии для всех случаев взаимодействий. Тем не менее при изучении целого ряда процессов в рамках классической термодинамики косвенным образом этот учет осуществляется. Известно, например, эмпирическое правило Трутона, согласно которому при испарении большинства жидкостей происходит одинаковое увеличение энтропии. Исключение из этого правила составляют жидкости, обладающие ассоциированными или поляризованными молекулами. Предлагаемая точка зрения проливает свет на этот факт. При испарении жидкости с поляризованными молекулами помимо возрастания энтропии, обусловленного переходом теплового движения в более неупорядоченное движение молекул пара, происходит дополнительная потеря информации за счет исчезновения упорядоченности движения поляризованных молекул. Поэтому ΔH поляризованной жидкости больше ΔH Трутона.

Таким образом, упорядоченность движения в той или иной степени присуща различным материальным системам, и на всех этапах развития увеличение упорядоченности микродвижения приводило к усложнению структуры систем. «Одним из основных принципом жизни,— пишет А. Сент-Дьердьи,— является «организация»; мы понимаем под этим, что при объединении двух вещей рождается нечто новое, качества которого не аддитивны и не могут быть выражены через качества составляющих его компонент. Это относится ко всей гамме форм организации, к объединению электронов и ядер, образующих атом, к соединению атомов в молекулы, аминокислот в пептиды, пептидов в белки, белков и нуклеиновых кислот в нуклеопротеиды и т. д.».

Причиной тенденции к упорядочиванию является одно общее свойство материальных систем: *степенью упорядоченности движения, существующего в той или иной системе, определяется влияние этой системы на упорядоченность движения взаимодействующих с ней систем.* Так, например, система, обладающая упорядоченной структурой кристалла, приобретает способность передавать «модели движения» молекулам жидкости, а живая клетка способна сохранять значительно более сложные «модели движения» в форме информации, имеющей определенный биологический код. Такая система приобретает возможности *управления,* благодаря которому она может сохранять качественную определенность и целостность в изменяющихся условиях внешней среды.

Наукой исследован целый ряд естественно протекающих сложных физико-химических процессов, сопровождающихся накоплением информации и соответствующим увеличением упорядоченности движения элементов взаимодействующих систем. Многие авторы рассматривают эти процессы как переходную ступень к тем сложным процессам самоорганизации, благодаря которым зарождается жизнь. К числу таких процессов относится, например, образование скоплений молекул органических веществ, упорядоченность которых регулируется гидрофильно-гидрофобной асимметрией каждой молекулы. В частности, наличие полярной группы и раздвоенного углеводородного «хвоста» у молекул липидов приводит к образованию сложноорганизованных слоистых (мембранных) структур. Другим примером могут служить процессы образования коацерватов, которые А. И. Опарин считает прообразом простейших одноклеточных структур, в которых зарождается жизнь.

Как показал М. Кальвин, упорядоченность движения, возникающая в подобных процессах, регулируется путем так называемого аутокатализа. При аутокатализе образующиеся в ходе цепной химической реакции вещества оказывают каталитическое стимулирующее воздействие на образование веществ в предыдущих звеньях этой цепи. С информационной точки зрения аутокатализ есть не что иное, как регулирование по обратной связи (недаром это явление называют также и *обратным катализом* — reflexive catalysis). Как установлено кибернетикой, обратная связь является наиболее эффективным средством регулирования движения всех упорядоченных систем.

Все рассмотренные нами примеры относятся к числу так называемых «антиэнтропийных» (или «негэнтропийных») процессов. В то же время второе начало термодинамики утверждает объективную необходимость возрастания энтропии как следствия равномерного распределения теплоты. Таким образом, развитие теории информации и появление меры, с помощью которой можно учитывать негэнтропию, со всей остротой ставит на повестку дня вопросы о том, как действует второе начало термодинамики в условиях взаимодействия энтропийных и негэнтропийных процессов.

Качественное отличие всех видов энергии от тепловой энергии обусловлено упорядоченностью движения, присущего электромагнитным, ядерным и гравитационным полям. Проявлением этой упорядоченности являются их волновые свойства, позволяющие априорно предугадывать характер движения, существующего в тот или иной момент времени в любой заданной точке пространства. Именно этими видами энергии обусловлено дальнейшее накопление информации в процессах концентрации материи, рассеянной в космосе, и включение ее в дальнейшие циклы развития (образование звезд и планет). Эти процессы являются «вещественными» аргументами против теории тепловой смерти Вселенной, основанной на рассмотрении свойств одной лишь тепловой энергии и процесса увеличения энтропии.

Отмеченное нами отличие тепловой энергии от всех других видов энергии нашло отражение еще в классической термодинамике и, в частности, в уравнении Гиббса— Гельмгольца:

$$\Delta U = \Delta F + T\Delta S,$$

где ΔU — полное приращение внутренней энергии физического тела; ΔF— приращение «свободной» энергии; $T\Delta S$ — приращение «связанной» энергии (приращение «связанной» энергии определяется произведением абсолютной температуры T на приращение физической энтропии ΔS, которая, в свою очередь, связана с рассматриваемой нами математической энтропией H соотношением $\Delta S = K\Delta H,$ где K — постоянная Больцмана.) Теперь становится очевидным, что выявленная классической термодинамикой необходимость раздельного рассмотрения двух форм внутренней энергии физических тел является отражением энтропийно-информационных соотношений термодинамических процессов: так называемая свободная энергия есть не что иное, как «информированная» («векторизованная») часть внутренней энергии (энергия всех видов полей), а обусловленная хаотичным микродвижением «связанная» энергия — это «энтропийная» часть внутренней энергии физических тел и систем.

Классическая термодинамика исследует процессы превращения энтропийной тепловой энергии в информированную (направленную) механическую энергию («полезную работу»), осуществляющиеся на макроскопическом уровне в специально созданных для этой цели тепловых машинах (цикл Карно). В природе аналогичные процессы превращения энтропийной энергии в информированную энергию поля осуществляются на микроскопическом уровне. **Информированной энергией в этом случае является дискретная (квантованная) и упорядоченная во времени и в пространстве (колебательная) энергия излучений, возникающих всюду, где имеется разность**

температур. Именно эта форма энергии (в частности, энергия излучений Солнца) является первоисточником всей информации, которую накапливает мир.

Информационные исследования описанных процессов несомненно позволят вскрыть более глубокую взаимосвязь механических и электродинамических свойств материи, корпускулярных и волновых свойств, законов взаимодействия и взаимных превращений элементарных частиц. Возникший и развивающийся благодаря теории информации новый взгляд на природу негэнтропийных процессов позволяет сделать обобщающий вывод о том, что **источником информации, вносящей упорядоченность, направленность в различные виды движения, является воздействие на это движение любых видов полей.** Так, например, в космогонических процессах эту роль выполняет гравитационное поле. Стоит лишь под действием сил инерции и гравитации из хаотически распыленных частиц возникнуть какому-то сгустку материи, как тут же он превращается в центр гравитационного притяжения, вносящий направленность и упорядоченность в движение всех расположенных в этом районе частиц. Соотношением сил инерции и сил притяжения регулируются размеры образующихся таким образом космических тел. **Когда возникающий в космосе сгусток материи достигает определенной плотности, дальнейшая упорядоченность космического движения регулируется действием ядерных полей**. Возникающие в плотных сгустках материи ядерные взрывы создают разность температур между космическими телами, а **разность температур или энергетических потенциалов**, как мы показали выше, **является первопричиной упорядоченного движения (излучения) и всех последующих стадий развития любых материальных систем**.

Можно предполагать, что наряду с законом всемирного тяготения в масштабах Вселенной действует также ***закон всемирного равновесия***, суть которого заключается в том, что **обесценивание энергии (рост энтропии), обусловленное рассеиванием *тепловой энергии,* компенсируется накоплением информации (уменьшением энтропии), обусловленным концентрацией массы, осуществляющейся под действием *гравитационных сил*.**

(Мысль об ограничении роста энтропии действием гравитационного поля высказывалась Я. П. Терлецким).

«Закон всемирного равновесия» может быть противопоставлен теории тепловой смерти Вселенной, вытекающей из одностороннего рассмотрения свойств тепловой энергии и не учитывающей упорядочивания космогонических процессов, обусловленного действием гравитационных и ядерных сил. **За счет взаимодействия**

тепловой, гравитационной и ядерной энергии осуществляется замкнутый космогонический цикл, изображенный на рис. 13.18.

Рис. 13.18.

В обозримые отрезки времени каждая локальная область Вселенной проходит одну из стадий цикла. Поскольку рассматриваемый цикл заключает в себе две противоположные тенденции (рост энтропии и рост информации), есть основания предполагать, что в масштабах Вселенной энтропия сохраняется постоянной, удовлетворяя условию $dH(x, y, z, t) = 0$ при неограниченно возрастающих (стремящихся к бесконечности) значениях x, y, z и t.

Приведенный нами краткий обзор процессов возникновения упорядоченного движения показывает, что существуют общие закономерности, действующие на всех уровнях организации материальных систем. С появлением сознания процесс накопления информации перестает быть стихийным. На этой стадии возникнет понятие *цели,* ради которой человек создает многообразные средства хранения, обработки и передачи информации. Научные теории и понятия, возникающие в процессе развития человеческой мысли, являются отраженными в нашем сознании «моделями движения» материальною мира.

14. Вычисление энтропии для частных случаев

В настоящем разделе излагаются методы подсчета энтропии многих случайных величин или случайного процесса в дискретном и непрерывном времени.

Особый интерес с принципиальной, а также прикладной точки зрения представляют стационарные случайные процессы и их теоретико-информационные характеристики, в частности, энтропия. Такие процессы являются относительно простыми объектами, особенно дискретный процесс, т. е. стационарный процесс с дискретными состояниями, протекающий в дискретном времени. Поэтому на его примере удобно иллюстрировать основные положения, и мы с него начнем.

Основное внимание будет уделено определению такой важной характеристики стационарного процесса, как удельная энтропия, т. е. энтропия, отнесенная к единице времени, к одному такту. Кроме того вводится энтропия Г конца отрезка. Она вместе с удельной энтропией H_1 определяет энтропию длинного отрезка длины T по приближенной формуле

$$H_{\text{Т}} \approx H_1 T + 2\,\Gamma,$$

которая тем точнее, чем больше T. Обе постоянные H_1, Г вычисляются для дискретного марковского процесса.

Обобщенное определение энтропии, данное в п. 14.3, позволяет применять это понятие как к непрерывно-значным случайным величинам, так и к тому случаю, когда число этих случайных величин континуально, т. е. к случайному процессу с непрерывным параметром (временем).

Ниже показано, что на случай непрерывного пространства значений и на случай непрерывного времени могут быть распространены многие результаты, относящиеся к дискретному процессу. Так, для стационарных процессов в непрерывном времени может быть введена удельная энтропия, отнесенная уже не к одному такту, а к единице времени, и энтропия конца отрезка. Энтропия случайного процесса на отрезке приближенно представляется в виде двух членов по аналогии с приведенной выше формулой.

Для нестационарных процессов с непрерывным временем, вместо постоянной удельной энтропии, нужно рассматривать плотность энтропии, которая, вообще говоря, непостоянна во времени.

Энтропия и ее плотность подсчитываются для различных важных случаев процессов с непрерывным временем: гауссовых процессов и диффузионных марковских процессов.

Проводимый здесь подсчет энтропии случайных процессов дает возможность подсчитывать для случайных процессов шенноновское количество информации.

14.1. Асимптотическая эквивалентность неравновероятных возможностей равновероятным

Идея о том, что общий случай неравновероятных возможностей (состояний) асимптотически сводится к случаю равновероятных, лежит в основе теории информации в отсутствие помех. Эта идея принадлежит Л. Больцману, который и вывел формулу для энтропии. К. Шеннон возродил эту идею и широко использовал для получения новых результатов.

При рассмотрении данного вопроса в настоящем пункте мы не будем стремиться к общности, поскольку эти результаты заведомо будут перекрыты более общими результатами в п. 14.2. Возьмем набор независимых реализаций $\eta=(\xi_1, \ldots, \xi_n)$ случайной величины $\xi=\xi_j$, принимающей одно из двух значений 1 или 0 с вероятностями $\mathbf{P}[\xi=1]=p<1/2$; $\mathbf{P}[\xi=0]=1-p=q$. Очевидно, что число различных таких наборов — реализаций равно 2^n. Пусть реализация η (обозначим ее η_{n_1}) имеет в своем составе n_1 единиц и $n-n_1=n_0$ нулей. Тогда ее вероятность равна

$$P(\eta_{n_1})=p^{n_1}q^{n-n_1}. \tag{14.1}$$

Конечно, эти вероятности при различных n_1 различны. Отношение $P(\eta_0)/P(\eta_n)=(q/p)^n$ самой большой вероятности к самой малой велико и сильно растет с ростом n. О какой же равновероятности можно говорить? Дело в том, что по закону больших чисел число единиц $n_1=\xi_1+\ldots+\xi_n$ имеет тенденцию принимать значения, близкие к своему среднему значению

$$\mathbf{M}n_1=\sum_{j=1}^{n}\mathbf{M}\xi_j=n\mathbf{M}\xi_j=np.$$

Найдем дисперсию $\mathbf{D}n_1=\mathbf{D}[\xi_1+\ldots+\xi_n]$ числа единиц. В силу независимости слагаемых имеем

$$\mathbf{D}n_1=n\mathbf{D}\xi=n[\mathbf{M}\xi^2-(\mathbf{M}\xi)^2],$$

причем

$$\mathbf{M}\xi^2 = \mathbf{M}\xi = p; \qquad \mathbf{M}\xi^2 - (\mathbf{M}\xi)^2 = p - p^2 = pq.$$

Следовательно,

$$\mathbf{D}n_1 = npq; \qquad \mathbf{D}(n_1/n) = pq/n. \tag{14.2}$$

Мы получили, таким образом, что среднее отклонение

$$\Delta n_1 = n_1 - np \sim \sqrt{pqn}$$

растет с ростом n, но медленнее, чем растет среднее значение np и длина всего возможного диапазона $0 \leq n_1 \leq n$. Типичное относительное отклонение $\Delta n_1/n_1$ убывает по закону $\Delta n_1/n_1 \sim \sqrt{q/np}$.

В пределах диапазона $|n_1 - pn| \sim \sqrt{pqn}$ неодинаковость вероятностей $P(\eta_{n_1})$ все-таки довольно велика:

$$\frac{P(\eta_{n_1})}{P(\eta_{n_1-\Delta n_1})} = \left(\frac{q}{p}\right)^{\Delta n_1} \approx \left(\frac{q}{p}\right)^{\sqrt{pqn}}$$

(как следует из (14.1)) и увеличивается с ростом n. Но это увеличение происходит относительно медленно, т. е. гораздо медленнее, чем увеличивается обратная величина самой вероятности. Соответствующее неравенство

$$\ln\frac{P(\eta_{n_1})}{P(\eta_{n_1-\Delta n_1})} \ll \ln\frac{1}{P(\eta_{n_1})}$$

при $\Delta n_1 \approx \sqrt{pqn}$ становится все более и более сильным с ростом n. Сформулируем сказанное точнее.

Теорема 1. *Все 2^n реализаций η можно разделить на два множества A_n и B_n так, что*

1) *суммарная вероятность реализаций одного множества (A_n) исчезает:*

$$\mathbf{P}(A_n) \to 0 \text{ при } n \to \infty; \tag{14.3}$$

2) *а реализации второго множества B_n становятся относительно равновероятными в следующем смысле:*

$$\left|\frac{\ln P(\eta) - \ln P(\eta')}{\ln P(\eta)}\right| \to 0, \quad \eta \in B_n;\ \eta' \in B_n. \tag{14.4}$$

Доказательство. Используя неравенство Чебышева, имеем

$$\mathbf{P}[\,|n_1 - pn| \geqslant \varepsilon] \leqslant \frac{\mathbf{D}n_1}{\varepsilon^2}.$$

Учитывая (14.2) и полагая $\varepsilon = n^{3/4}$, находим отсюда

$$\mathbf{P}[\,|n_1 - pn| \geqslant n^{3/4}] \leqslant pqn^{-1/2}. \tag{14.5}$$

Реализации η_{n_1} , для которых $| n_1 — pn | \geq n^{3/4}$, отнесем к множеству A_n, а остальные — к множеству B_n. Тогда первая часть (14.5) будет $P(A_n)$ и предельный переход $n\to 0$ в (14.5) докажет (14.3).

Для реализаций второго множества справедливо неравенство $pn — n^{3/4}< n_1< pn + n^{3/4}$, из которого, учитывая (14.1), получаем

$$-n(p\ln p+q\ln q)-n^{3/4}\ln\frac{q}{p}<-\ln P(\eta_{n_1})<$$

$$<-n(p\ln p+q\ln q)+n^{3/4}\ln\frac{q}{p}. \tag{14.6}$$

Отсюда имеем

$$|\ln P(\eta)-\ln P(\eta')|<2n^{3/4}\ln\frac{q}{p} \quad \text{при } \eta\in B_n,\ \eta'\in B_n$$

и

$$|\ln P(\eta)|>-n(p\ln p+q\ln q)-n^{3/4}\ln\frac{q}{p}.$$

Следовательно,

$$\left|\frac{\ln P(\eta)-\ln P(\eta')}{\ln P(\eta)}\right|<\frac{2n^{-1/4}\ln\frac{q}{p}}{-p\ln p-q\ln q-n^{-1/4}\ln\frac{q}{p}}.$$

Чтобы получить (14.4), здесь остается совершить предельный переход $n\to\infty$. Доказательство закончено.

Неравенство (14.6) позволяет, кроме того, оценить число элементов множества B_n.

Теорема 2. *Пусть B_n — множество, описанное в теореме* 1. *Число M его элементов таково, что*

$$\frac{\ln M}{n}\to-p\ln p-q\ln q\equiv H_{\xi_1} \quad \text{при} \quad n\to\infty. \tag{14.7}$$

Эту теорему можно сформулировать также в следующей форме. **Теорема 3**. *Если считать реализации множества A_n имеющими нулевую вероятность, а реализации множества B_n равновероятными и вычислять энтропию H_η по простой формуле H_η =* ln *M, то удельная энтропия H_η /n в пределе будет совпадать с энтропией, вычисленной по формуле*

$$\lim_{n\to\infty}\frac{\widetilde{H}_\eta}{n}=-p\ln p-q\ln q. \tag{14.8}$$

Доказательство. Согласно (14.5) сумма

$$\sum_{\eta \in B_n} P(\eta) = P(B_n)$$

оценивается неравенством

$$P(B_n) = 1 - P(A_n) \geqslant 1 - \frac{pq}{\sqrt{n}}\ .$$

Указанная сумма распространяется на элементы множества B_n и имеет M членов. Вследствие (14.6) каждый член можно оценить сверху:

$$P(\eta) < \exp\left[n(p\ln p + q\ln q) + n^{3/4}\ln\frac{q}{p}\right].$$

Следовательно, число членов не может быть меньше определенного выражения

$$M \geqslant \frac{P(B_n)}{\max P(\eta)} > \left(1 - \frac{pq}{\sqrt{n}}\right)\exp\left[-n(p\ln p + q\ln q) - -n^{3/4}\ln\frac{q}{p}\right]. \tag{14.9}$$

С другой стороны, $\sum_{B_n} P(\eta) \leqslant 1$ и в силу (14.6)

$$P(\eta) > \exp\left[n(p\ln p + q\ln q) - n^{3/4}\ln\frac{q}{p}\right],$$

так что

$$M \leqslant \frac{P(B_n)}{\min P(\eta)} \leqslant \frac{1}{\min P(\eta)} < \exp\left[-n(p\ln p + q\ln q) + +n^{3/4}\ln\frac{q}{p}\right]. \tag{14.10}$$

Логарифмируя полученные неравенства (14.9), (14.10), имеем

$$H_{\xi_1} - n^{-1/4}\ln\frac{q}{p} + \ln\left(1 - \frac{pq}{\sqrt{n}}\right) < \frac{\ln M}{n} < H_{\xi_1} + +n^{-1/4}\ln\frac{q}{p}\ .$$

Предельным переходом $n \to \infty$ доказываем требуемые соотношения (14.7), (14.8).

В случае, когда ξ_1 принимает одно из m значений, имеется m^n различных реализаций процесса $\eta=(\xi_1, ..., \xi_n)$ с независимыми значениями. Согласно вышеизложенному из них заслуживают внимания лишь $M = e^{nH\xi 1}$ реализаций, которые можно считать равновероятными. Когда $P(\xi_1) = 1/m$, $H_{\xi 1} = \ln m$, указанные два числа равны один другому, в противном случае, когда $H_{\xi 1} < \ln m$, доля $e^{nH\xi 1}/m^n$ заслуживающих внимания реализаций неограниченно уменьшается с ростом n. Следовательно, подавляющее большинство реализаций при этом является несущественным и его можно отбросить. Этот факт лежит в основе теории кодирования .

Рассмотренный выше факт асимптотической равновероятности имеет место и при более общих предположениях в случае зргодических стационарных процессов. Указанные равновероятные реализации Больцман называл «микросостояниями», противопоставляя их «макросостояниям», образованным ансамблем «микросостояний».

14.2. Асимптотическая равновероятность и энтропийная устойчивость

1. Изложенные в предыдущем пункте идеи, касающиеся асимптотической эквивалентности неравновероятных возможностей (сообщений) равновероятным, могут быть распространены на значительно более общий класс случайных последовательностей и процессов. Необязательно, чтобы случайные величины ξ_j, образующие отрезок последовательности $\eta^n = (\xi_1, ..., \xi_n)$, принимали лишь одно из двух значений. Необязательно, чтобы они имели одинаковый закон распределения $P(\xi_j)$. Необязательно, наконец, чтобы они были статистически независимыми и не обязательно даже, чтобы η^n была последовательностью $(\xi_1, ..., \xi_n)$. Что же является обязательным для описанной асимптотической эквивалентности?

Дать общую формулировку свойства асимптотической эквивалентности неравновероятных возможностей равновероятным помогает понятие энтропийной устойчивости семейства случайных величин.

Семейство случайных величин $\{\eta^n\}$ является *энтропийно устойчивым,* если отношение $H(\eta^n)/H_{\eta^n}$ при $n\to\infty$ сходится по вероятности к единице. Это значит, что каковы бы ни были $\varepsilon > 0$, $\eta > 0$, найдется такое $N(\varepsilon, \eta)$, что будет выполнено неравенство

$$\mathbf{P}\{|H(\eta^n)/H_{\eta^n} - 1| \geqslant \varepsilon\} < \eta \tag{14.11}$$

при любом $n \geq N(\varepsilon, \eta)$.

В приведенном выше определении подразумевается, что все

$0< H_{\eta^n} <\infty$ и H_{η^n} не убывают с ростом n. Обычно $H_{\eta^n} \to\infty$.

Факт асимптотической равновероятности можно сформулировать при помощи понятия энтропийной устойчивости в виде следующей общей теоремы.

Теорема 4. *Если семейство случайных величин $\{\eta^n\}$ является энтропийно устойчивым, то множество реализаций каждой случайной величины можно разбить на два подмножества A_n и B_n таким образом, что*

1) *суммарная вероятность реализаций подмножества A_n исчезает:*

$$P\{A_n) \to 0 \text{ при } n\to\infty; \tag{14.12}$$

2) *реализации второго подмножества B_n становятся относительно равновероятными в смысле соотношения*

$$\left|\frac{\ln P(\eta)-\ln P(\eta')}{\ln P(\eta)}\right| \to 0 \text{ при } n\to\infty,\ \eta\in B_n, \eta'\in B_n; \tag{14.13}$$

3) *число M_n реализаций множества B_n связано с энтропией H_{η^n} соотношением*

$$\ln M_n/H_{\eta^n} \to 1 \text{ при } n\to\infty. \tag{14.14}$$

Доказательство. Полагая, например, $\varepsilon = 1/m$, $\eta = 1/m$, из (14.11) имеем

$$\mathbf{P}\{|H(\eta^n)/H_{\eta^n}-1|\geqslant 1/m\}<1/m \tag{14.15}$$

при $n\geq N(1/m,\ 1/m)$. Пусть m увеличивается, пробегая последовательные целые значения. Множества A_n определяем как множества реализаций, удовлетворяющих неравенству

$$|H(\eta^n)/H_{\eta^n}-1|\geqslant 1/m \tag{14.16}$$

при $N(1/m, 1/m) \leq n< N(1/(m+1), 1/(m+1))$. При таком определении, очевидно, свойство (14.12) выполняется в силу (14.15). Для дополнительного множества B_n имеем

$$(1-1/m)H_{\eta^n}< -\ln P(\eta^n)<(1+1/m)\,H_{\eta^n}\quad (\eta^n\in B_n), \tag{14.17}$$

откуда получаем

$$\left|\frac{\ln P(\eta^n)-\ln P(\eta'^n)}{-\ln P(\eta^n)}\right|\leqslant\frac{2H_{\eta^n}/m}{-\ln P(\eta^n)}\leqslant\frac{2/m}{1-1/m}$$

при $\eta^n\in B_n,\ \eta'^n\in B_n,$ что доказывает свойство (14.13) (сходимость $n = N(1/m, 1/m)$ вызывает сходимость $m\to\infty$).

Вероятности $P\ (\eta^n)$ всех реализаций множества B_n согласно (14.17) лежат в диапазоне

$$e^{-(1+1/m)H_{\eta^n}} < P(\eta^n) < e^{-(1-1/m)H_{\eta^n}}.$$

В то же время суммарная вероятность $P\ (B_n) = \sum_{B_n} P\ (\eta^n)$ заключена между 1 — 1/m и 1. Отсюда получаем следующий диапазон для числа членов:

$$(1-1/m)\,e^{(1-1/m)H_{\eta^n}} < M_n < e^{(1+1/m)H_{\eta^n}}$$

(слева наименьшее число делится на наибольшее, а справа — наибольшее на наименьшее). Поэтому

$$1-\frac{1}{m}+\frac{\ln(1-1/m)}{H_{\eta^n}} < \frac{\ln M_n}{H_{\eta^n}} < 1+\frac{1}{m},$$

откуда вытекает (14.14). (Член $\ln(1-1/m)/H_{\eta^n}$ стремится к нулю вследствие неубывания энтропии H_{η^n}) . Доказательство закончено.

Свойство энтропийной устойчивости, играющее согласно теореме 4 большую роль, удобно проверять для различных примеров, вычисляя дисперсию

$$\mathbf{D}H(\eta^n) = \mathbf{M}\,H^2(\eta^n) - H^2_{\eta^n}$$

случайной энтропии. Если эта дисперсия не слишком быстро растет с ростом n, то применением неравенства Чебышева можно доказать (14.11), т. е. энтропийную устойчивость. Сформулируем три относящиеся к этому вопросу теоремы.

Теорема 5. *Если существует равный нулю предел*

$$\lim_{n\to\infty}\frac{\mathbf{D}H(\eta^n)}{H^2_{\eta^n}} = 0, \tag{14.18}$$

то семейство случайных величин $\{\eta^n\}$ *является энтропийно устойчивым.*

Доказательство. Согласно неравенству Чебышева для любой случайной величины с конечной дисперсией при любом ε> 0 имеем

$$\mathbf{P}\{|\xi-\mathbf{M}\,\xi|\geqslant\varepsilon\}\leqslant\mathbf{D}\xi/\varepsilon^2.$$

Полагая здесь $\xi = H\ (\eta_{\eta^n})/\ H_{\eta^n}$ и учитывая (14.18), получаем

$$\mathbf{P}\left\{\left|\frac{H(\eta^n)}{H_{\eta^n}}-1\right|\geqslant\varepsilon\right\}\to 0 \quad \text{при} \quad n\to 0$$

при любом ε. Отсюда вытекает (14.11), т. е. энтропийная устойчивость.

Теорема 6. *Если энтропия* H_{η^n} *неограниченно возрастает и существует ограниченный верхний предел*

$$\limsup_{n\to\infty} \frac{\mathbf{D}H(\eta^n)}{H_{\eta^n}} < C, \tag{14.19}$$

то семейство случайных величин энтропийно устойчиво.

Для доказательства нужно лишь учесть, что величина $\mathbf{D}H(\eta^n)/H^2_{\eta^n}$, которую в силу (14.19) можно оценить как

$$\frac{\mathbf{D}H(\eta^n)}{H^2_{\eta^n}} \leqslant \frac{C}{H_{\eta^n}} + \varepsilon \quad (\text{где} \quad n > N(\varepsilon)),$$

стремится к нулю вследствие возрастания H_{η^n} (и произвольности ε) так что к данному случаю можно применить теорему 5.

Во многих важных для практики случаях существуют конечные пределы

$$H_1 = \lim_{n\to\infty} \frac{1}{n} H_{\eta^n}; \qquad D_1 = \lim_{n\to\infty} \frac{1}{n} \mathbf{D}H(\eta^n), \tag{14.20}$$

которые можно назвать удельной энтропией и удельной дисперсией. Для их вычисления разработан ряд более или менее общих методов. Согласно приводимой ниже теореме конечность этих пределов гарантирует энтропийную устойчивость.

Теорема 7. *Если пределы* (14.20) *существуют и конечны и первый из них отличен от нуля, то семейство случайных величин энтропийно устойчиво.*

Для доказательства учтем, что из (14.20) следуют соотношения

$$H_{\eta^n} = H_1 n + o(n); \tag{14.21}$$

$$\mathbf{D}H(\eta^n) = D_1 n + o(n). \tag{14.22}$$

Здесь, как обычно, *o(n)* обозначает, что $o(n)/n \to 0$. Поскольку $H_1>0$, энтропия H_{η^n} неограниченно нарастает. Поделив выражение (14.21) на (14.22), получим конечный предел

$$\lim_{n\to\infty} \frac{\mathbf{D}H(\eta^n)}{H_{\eta^n}} = \lim_{n\to\infty} \frac{D_1 + o(1)}{H_1 + o(1)} = \frac{D_1}{H_1}.$$

Тем самым выполнены условия теоремы 6, что доказывает энтропийную устойчивость.

Пример. Пусть задана бесконечная последовательность $\{\xi_i\}$ дискретных случайных величин, статически независимых, но по разному распределенных. Предположим, что энтропии H_{ξ_j} всех этих величин конечны и равномерно ограничены снизу.

$$H_{\xi_j} > C_1, \tag{14.23}$$

а дисперсии равномерно ограничены сверху:

$$\mathbf{D}H(\xi_j) < C_2. \tag{14.24}$$

Случайные величины η^n определяем как набор n первых элементов последовательности $\eta^n = \{ \xi_1, ..., \xi_n\}$. Тогда вследствие (14.20), (14.24) и статистической независимости, т. е. аддативности энтропии, будем иметь

$$H_{\eta^n} > C_1 n; \qquad \mathbf{D}H(\eta^n) < C_2 n.$$

В этом случае условия теоремы 6 являются выполненными и из нее вытекает энтропийная устойчивость семейства $\{\eta^n\}$.

Другие более сложные примеры энтропийно устойчивых случайных величин, не распадающихся на статистически независимые величины, будут рассмотрены в дальнейшем.

Понятие энтропийной устойчивости можно, правда менее строго, применять не к последовательности случайных величин $\{\eta^n\}$, а к одной случайной величине η. Тогда под энтропийно устойчивой случайной величиной η нужно понимать такую величину, для которой при некотором малом значении ε вероятность

$$\mathbf{P}\left\{\left|\frac{H(\eta)}{H_\eta} - 1\right| < \varepsilon\right\}$$

достаточно близка к единице, т. е. $H(\eta)/H_\eta$ достаточно близко к единице.

В дальнейшем будут введены другие понятия: понятие информационной устойчивости, канонической устойчивости и пр., которые во многих отношениях напоминают энтропийную устойчивость.

2. Для получения ряда результатов, связанных с энтропийно-устойчивыми случайными величинами, иногда удобно рассматривать характеристический потенциал μ_0 (α) энтропии H (η), определяемый формулой

$$e^{\mu_0(\alpha)} = \sum_\eta e^{\alpha H(\eta)} P(\eta) = \sum_\eta P^{1-\alpha}(\eta), \tag{14.25}$$

При помощи этого потенциала удобно исследовать быстроту сходимости (14.12)—(14.14) в теореме 4. Этот вопрос затрагивается в следующей теореме.

Теорема 8. *Пусть потенциал* (14.25) *определен и дифференцируем в отрезке* $s_1 < \alpha < s_2$ ($s_1 < 0$; $s_2 > 0$) *и пусть уравнение*

$$\frac{d\mu_0}{d\alpha}(s) = (1+\varepsilon) H_\eta \qquad (\varepsilon > 0) \tag{14.26}$$

имеет корень $s \in [0,\ s_2]$. *Тогда подмножество А реализаций* η, *определенное условием*

$$\frac{H(\eta)}{H_\eta} - 1 > \varepsilon, \tag{14.27}$$

имеет вероятность

$$\mathbf{P}(A) \leqslant e^{-s\mu'_0(s) + \mu_0(s)}. \tag{14.28}$$

Остальные реализации, составляющие дополнительное подмножество В, имеют вероятности, связанные соотношением

$$\left| \frac{\ln P(\eta) - \ln P(\eta')}{\ln P(\eta)} \right| < \frac{2\varepsilon}{1-\varepsilon} \quad (\eta,\ \eta' \in B), \tag{14.29}$$

причем число М этих реализаций удовлетворяет неравенству

$$1 - \varepsilon + \frac{1}{H_\eta} \ln\left[1 - e^{-s\mu'_0(s) + \mu_0(s)}\right] < \frac{\ln M}{H_\eta} < 1 + \varepsilon. \tag{14.30}$$

Доказательство во многом аналогично доказательству теоремы 4. Соотношение (14.29) вытекает из (14.27). Неравенство (14.27) эквивалентно неравенству

$$e^{-(1+\varepsilon) H_\eta} < P(\eta) < e^{-(1-\varepsilon) H_\eta}.$$

Учитывая его и равенство

$$\sum_B P(\eta) - 1 - \mathbf{P}(A),$$

находим, что число членов в этой сумме, т. е. число реализаций в *В,* удовлетворяет неравенству

$$[1 - \mathbf{P}(A)]\, e^{(1-\varepsilon) H_\eta} < M < [1 - \mathbf{P}(A)]\, e^{(1+\varepsilon) H_\eta}.$$

Следовательно,

$$(1-\varepsilon) H_\eta + \ln[1 - \mathbf{P}(A)] < \ln M < (1+\varepsilon) H_\eta + \ln[1 - \mathbf{P}(A)].$$

Отсюда вытекает (14.30), если учесть отрицательность ln [1 — **P** *(A*)] и неравенство (14.28). Итак, для завершения доказательства теоремы остается обосновать (14.28). Это соотношение мы получим, при ξ = η, *B* (ξ) = *H* (η)). Доказательство закончено.

Из формул, приведенных в предыдущей теореме, можно получить ряд простых приближенных соотношений, если использовать условие малости ε. При ε = 0 имеется нулевой корень s = 0 уравнения (14.26), так как $\frac{d\mu}{d\alpha}(0) = H_\eta$. При малых ε значение s мало, и правую часть уравнения (14.26) можно разложить в ряд Маклорена

$$\mu_0'(s) = \mu_0'(0) + \mu_0''(0)\,s + \ldots,$$

так что корень уравнения будет иметь вид

$$s = \frac{\varepsilon H_\eta}{\mu_0''(0)} + O(\varepsilon^2). \tag{14.31}$$

Разложим далее в ряд выражение, стоящее в экспоненте (14.28):

$$s\mu_0'(s) - \mu_0(s) = \frac{1}{2}\mu_0''(0)\,s^2 + O(s^3).$$

Подставляя сюда (14.31), получаем

$$\mathbf{P}(A) \leqslant \exp\left\{-\frac{\varepsilon^2 H_\eta^2}{2\mu_0''(0)}\right\}[1 + O(\varepsilon^3)]. \tag{14.32}$$

Здесь $\mu''(0) > 0$ — положительная величина, в силу общих свойств характеристического потенциала, равная дисперсии

$$\mu_0''(0) = \mathbf{D}H(\eta). \tag{14.33}$$

Мы видим, что характеристический потенциал энтропии помогает исследовать вопросы, связанные с энтропийной устойчивостью.

14.3. Определение энтропии непрерывной случайной величины

До сих пор предполагалось, что случайная величина ξ, энтропия H_ξ которой изучалась, может принимать значения из некоторого дискретного пространства, состоящего из конечного или счетного числа элементов, например, сообщений, символов и т. п. Между тем в теории информации большое распространение имеют также непрерывные величины, т. е. величины (скалярные или векторные), которые могут принимать значения из непрерывного пространства X, чаще всего пространства действительных чисел. При этом случайная величина ξ описывается плотностью распределения вероятностей p (ξ). задающей вероятность

$$\Delta P = \int_{\xi \in \Delta X} p(\xi)\,d\xi \approx p(A)\,\Delta V \qquad (A \in \Delta X)$$

попадания ξ в область ΔX указанного пространства X, имеющую объем ΔV ($d\,\xi = dV$ — дифференциал объема).

Как определить энтропию H_ξ такой случайной величины? Один из возможных формальных путей таков. В формуле

$$H_\xi = -\sum_\xi P(\xi) \ln P(\xi) = -\mathbf{M} \ln P(\xi), \tag{14.34}$$

пригодной для дискретной величины, вероятности P (ξ) под знаком логарифма формально заменяются на плотность вероятности и берется, следовательно, выражение

$$H_\xi = \mathbf{M} \ln p(\xi) = -\int_X p(\xi) \ln p(\xi)\, d\xi. \tag{14.35}$$

Такой способ определения энтропии не очень обоснован. Остается неясным, как определять энтропию в комбинированном случае, когда в непрерывном пространстве, кроме непрерывного распределения, имеется еще концентрация вероятности в отдельных точках, т. е. плотность вероятности содержит еще дельта-образные особенности. Энтропия (14.35) обладает также тем недостатком, что является неинвариантной, т. е. меняется при невырожденном преобразовании переменных $\eta = f(\xi)$ в отличие от энтропии (14.34), которая при этом остается инвариантной.

Поэтому целесообразно дать несколько другое определение энтропии непрерывной случайной величины. Подойдем к этому определению исходя из формулы (14.35). Будем предполагать, что ξ — дискретная случайная величина, вероятности которой сосредоточены в точках A_i непрерывного пространства X:

$$P(A_i) = w_i.$$

Это означает, что плотность распределения имеет вид

$$p(\xi) = \sum_i w_i \delta(\xi - A_i), \qquad \xi \in X. \tag{14.36}$$

Используя формулу (14.34), в этом случае имеем

$$H_\xi = -\sum_i P(A_i) \ln P(A_i). \tag{14.37}$$

Пусть далее точки A_i расположены в пространстве X довольно плотно, так что даже в сравнительно малой области ΔX, имеющей объем ΔV, располагается довольно большое число ΔN их. Область ΔX предполагается малой в том смысле, что внутри ее вероятности $P(A_i)$ приблизительно равны

$$P(A_i) \approx \frac{\Delta P}{\Delta N} \quad \text{при} \quad A_i \in \Delta X. \tag{14.38}$$

Тогда для суммы по точкам, лежащим внутри ΔX, будем иметь

$$-\sum_{A_i \in \Delta X} P(A_i) \ln P(A_i) \approx -\Delta P \ln \frac{\Delta P}{\Delta N}.$$

Суммируя по всем таким областям ΔX, увидим, что энтропия (14.37) запишется в виде

$$-\sum_{A_i \in \Delta X} P(A_i) \ln P(A_i) \approx -\Delta P \ln \frac{\Delta P}{\Delta N}. \tag{14.39}$$

Если ввести меру v_0 (ξ), указывающую плотность точек A_i интегрированием которой вычисляется число точек

$$\Delta N = \int_{\xi \in \Delta X} \nu_0(\xi)\, d\xi$$

внутри любой области ΔX, то энтропию (14.37) можно записать

$$H_\xi = -\int_X p(\xi) \ln \frac{p(\xi)}{\nu_0(\xi)}\, d\xi. \tag{14.40}$$

Плотности распределения (14.36), очевидно, соответствует дельта-образная плотность v_0 (ξ) вида

$$\nu_0(\xi) = \sum_i \delta(\xi - A_i).$$

От этих дельта-образных плотностей можно перейти к сглаженным плотностям

$$\tilde{p}(\xi) = \int K(\xi - \xi')\, p(\xi')\, d\xi',$$
$$\tilde{\nu}_0(\xi) = \int K(\xi - \xi')\, \nu_0(\xi')\, d\xi'. \tag{14.41}$$

При этом, если «радиус сглаживания» r_0, соответствующий ширине функции $K(x)$, относительно невелик (нужно, чтобы на таких расстояниях вероятности $P(A_i)$ не успевали существенно измениться), то сглаживание мало повлияет на отношение плотностей:

$$\frac{\tilde{p}(\xi)}{\tilde{\nu}_0(\xi)} \approx \frac{p(\xi)}{\nu_0(\xi)}.$$

Поэтому из (14.40) будем иметь

$$H_\xi \approx -\int \tilde{p}(\xi) \ln \frac{\tilde{p}(\xi)}{\tilde{\nu}_0(\xi)}\, d\xi. \tag{14.42}$$

Эту формулу можно получить также из (14.39), так как $\Delta P \approx \Delta\, \tilde{p}\, AV$, $\Delta N \approx \tilde{V}_0 \Delta V$, когда «радиус» r_0 много меньше размеров областей ΔX. Но если «радиус сглаживания» r_0 значительно превосходит среднее расстояние между точками A_i, то сглаженные функции (14.41) будут иметь простой (гладкий) вид, который предполагался, скажем, в формуле (14.35).

Отбросив значки ~ в (14.42) у обеих плотностей, мы, следовательно, получим в качестве формулы, определяющей энтропию H_ξ, вместо (14.35) формулу

$$H_\xi = -\int_X p(\xi) \ln \frac{p(\xi)}{\nu_0(\xi)} d\xi. \tag{14.43}$$

Здесь ν_0 (ξ) — некоторая вспомогательная плотность, которая предполагается заданной. Она необязательно нормирована на единицу. Более того, в соответствии с приведенной выше интерпретацией энтропии (14.43) как частного случая энтропии (14.34) нормировочный интеграл

$$\int_X \nu_0(\xi)\, d\xi = N \tag{14.44}$$

предполагается достаточно большим числом ($N \gg 1$), интерпретируемым как общее число дискретных точек A_i в которых сконцентрированы вероятности $P(A_i)$. Такая интерпретация, однако, необязательна, и можно, в частности, брать N=1. Если ввести нормированную плотность

$$q(\xi) = \nu_0(\xi)/N \tag{14.45}$$

(что можно сделать, когда N конечно), то, очевидно, из формулы (14.43) будем иметь

$$H_\xi = \ln N - \int p(\xi) \ln \frac{p(\xi)}{q(\xi)} d\xi. \tag{14.46}$$

Определение (14.43) энтропии H_ξ может быть обобщено и на комбинированный случай, и на общий случай абстрактной случайной величины. Последняя считается заданной, если фиксировано некоторое пространство X точек ξ и борелевское поле F (σ-алгебра) его подмножеств. При этом говорят, что задано измеримое пространство (X, F). На этом поле далее определена вероятностная мера $P(A)$ $(A \in F)$, т. е. такая мера, что $P(X)$ = 1. Для определения энтропии мы требуем, чтобы на измеримом пространстве (X, F) была задана также вспомогательная мера $\nu(A)$ такая, что мера P абсолютно непрерывна относительно ν.

Мера P называется абсолютно непрерывной относительно меры ν, если для каждого множества A из F, для которого $\nu(A)$ = 0, имеет место равенство $P(A)$ = 0. Согласно теореме Радона— Никодима из условия абсолютной непрерывности меры P относительно меры ν вытекает существование F-измеримой функции $f(x)$, обозначаемой $dP/d\nu$ и называемой производной Радона—Никодима, которая обобщает понятие плотности распределения. Она определена для всех точек пространства *X,* за исключением, может быть, некоторого подмножества Λ, для которого $v(\Lambda) = 0$, а значит и $P(\Lambda) = 0$.

Итак, если условие абсолютной непрерывности выполнено, то энтропия H_ξ определяется при помощи производной Радона—Никодима по формуле

$$H_\xi = -\int\limits_{X-\Lambda-\Lambda_0} \ln \frac{dP}{d\nu}(\xi)\, P(d\xi). \tag{14.47}$$

Подмножество Λ, для которого функция dP/dv не определена, не влияет на результат интегрирования, так как ему соответствуют нулевые меры $P(\Lambda)=v(\Lambda)=0$. Также несущественным является то подмножество Λ_0, в котором функция dP/dv определена, но равна нулю, так как для него

$$P(\Lambda_0) = \int\limits_{\Lambda_0} \frac{dP}{d\nu}\, \nu(d\xi) = \int\limits_{\Lambda_0} 0 \cdot \nu(d\xi) = 0,$$

даже если $v(\Lambda_0) \neq 0$. Поэтому некоторая неопределенность функции $f(\xi) = dP/dv$ и бесконечные значения $\ln f(\xi)$ в точках, где $f(\xi) = 0$, не мешают определению (1.6.13) энтропии H_ξ в случае абсолютной непрерывности P относительно v. Величина

$$H(\xi) = -\ln \frac{dP}{d\nu}(\xi) \tag{14.48}$$

при этом играет роль случайной энтропии. Она определена почти всюду в *X,* т. е. во всем пространстве, за исключением, может быть, множества $\Lambda + \Lambda_0$ нулевой вероятности *P.*

По аналогии с (14.45), если $N = v(X) \to \infty$, можно ввести вместо $v(A)$ нормированную, т. е. вероятностную меру

$$Q(A) = \nu(A)/N, \qquad A \in F, \tag{14.49}$$

и преобразовать (14.47) к виду

$$H_\xi = \ln N - \int \ln \frac{dP}{dQ}(\xi)\, P(d\xi), \tag{14.50}$$

аналогичному (14.46). Величина

$$H_\xi^{P/Q} = \int \ln \frac{dP}{dQ}(\xi)\, P(d\xi) \tag{14.51}$$

является неотрицательной. Вследствие указанной неотрицательности величину (14.51) можно называть энтропией распределения вероятности *P* относительно распределения вероятности Q.

Данное в настоящем пункте определение энтропии (14.44), (14.47) позволяет рассматривать энтропию (14.35) как частный случай этого общего определения. Именно формула (14.35) есть энтропия (14.47) для того случая, когда мере v соответствует равномерная единичная плотность $v_0(\xi) = dv/d\xi = 1$.

Нужно отметить, что энтропия (14.51) вероятностной меры P относительно вероятностной меры Q может быть использована как показатель степени различия мер P и Q . Этому благоприятствует то обстоятельство, что она обращается в нуль для совпадающих мер $P(\bullet) = Q(\bullet)$ и положительна для несовпадающих.

Другим показателем различия мер, обладающим этими же свойствами, может служить «расстояние», определяемое формулой

$$s(P, Q) = \inf \int_P^Q ds, \tag{14.52}$$

причем

$$ds^2 = 2H^{P/P \pm \delta P} = 2\int \ln \frac{P(d\zeta)}{P(d\zeta) \pm \delta P(d\zeta)} P(d\zeta). \tag{14.53}$$

Произведя разложение функции —ln (1 ± $\delta P/P$) по $\delta P/P$, нетрудно убедиться, что эта метрика может быть задана также эквивалентной формулой

$$ds^2 = \int \frac{[\delta P(d\zeta)]^2}{P(d\zeta)} = \int [\delta \ln P(d\zeta)]^2 P(d\zeta). \tag{14.54}$$

Соединим точки P и Q «линией» — семейством точек P_λ, зависящих от параметра λ так, что $P_0 = P$, $P_1 = Q$. Тогда (14.52) можно записать

$$s(P, Q) = \inf \int_0^1 \frac{ds}{d\lambda} d\lambda, \tag{14.55}$$

где в силу (14.54)

$$\left(\frac{ds}{d\lambda}\right)^2 = \int \left[\frac{\partial \ln P_\lambda(d\zeta)}{\partial \lambda}\right]^2 P_\lambda(d\zeta). \tag{14.56}$$

Здесь и ниже предполагаются выполненными условия дифференцируемости. Из (14.56) вытекает, что

$$\left(\frac{ds}{d\lambda}\right)^2 = -\int \frac{\partial^2 \ln P_\lambda(d\zeta)}{\partial^2 \lambda} P_\lambda(d\zeta). \tag{14.57}$$

В самом деле, разность этих выражений

$$\int \left\{ \left[\frac{\partial \ln P_\lambda (d\zeta)}{\partial \lambda} \right]^2 + \frac{\partial^2 \ln P_\lambda (d\zeta)}{\partial^2 \lambda} \right\} P_\lambda (d\zeta) = \\ = \frac{\partial}{\partial \lambda} \int \frac{\partial \ln P_\lambda (d\zeta)}{\partial \lambda} P_\lambda (d\zeta)$$

равна нулю вследствие того, что тождественно исчезает выражение

$$\int \frac{\partial \ln P_\lambda (d\zeta)}{\partial \lambda} P_\lambda (d\zeta) = \frac{\partial}{\partial \lambda} \int P_\lambda (d\zeta) = \frac{\partial}{\partial \lambda} 1.$$

Из указанного определения видно, что при сближении точек P и Q энтропии $H^{p/Q}$, $H^{Q/p}$ и величина $\frac{1}{2} s^2 (P, Q)$ перестают различаться. Можно сформулировать, однако, теорему, связывающую указанные величины не только в случае близких точек.

Теорема 9. *Квадрат «расстояния»* (14.52) *ограничен сверху суммой энтропии:*

$$s^2 (P, Q) \leqslant H^{P/Q} + H^{Q/P}. \tag{14.58}$$

Доказательство. Соединим точки P и Q линией

$$P_\lambda (d\zeta) = e^{-\Gamma(\lambda)} [P (d\zeta)]^{1-\lambda} Q (d\zeta)^\lambda \\ (\Gamma (\lambda) = \ln \int [P (d\zeta)]^{1-\lambda} [Q (d\zeta)]^\lambda). \tag{14.59}$$

Для нее получаем

$$-\frac{\partial^2 \ln P_\lambda (d\zeta)}{\partial^2 \lambda} = \frac{d^2 \Gamma (\lambda)}{d\lambda^2}.$$

Следовательно, в силу (14.57)

$$\left(\frac{ds}{d\lambda} \right)^2 = \frac{d^2 \Gamma}{d\lambda^2},$$

также

$$-\frac{\partial^2 \ln P_\lambda (d\zeta)}{\partial^2 \lambda} = \left(\frac{ds}{d\lambda} \right)^2. \tag{14.60}$$

Далее нетрудно убедиться, что выражение (14.51) можно записать в такой интегральной форме

$$H^{P/Q} = -\int_0^1 d\lambda \int \frac{\partial \ln P_\lambda (d\zeta)}{\partial \lambda} P_0 (d\zeta) = \\ = -\int_0^1 d\lambda \int_0^\lambda d\lambda' \int \frac{\partial^2 \ln P_{\lambda'} (d\zeta)}{\partial \lambda'^2} P_0 (d\zeta).$$

Здесь учтено, что

$$\int [\partial (\ln P_\lambda (d\zeta))/\partial\lambda]\, P_0 (d\zeta) = 0 \text{ при } \lambda = 0.$$

Принимая во внимание (14.60), отсюда находим

$$H^{P/Q} = \int_0^1 d\lambda \int_0^\lambda \left(\frac{ds}{d\lambda}\right)^2 d\lambda'_t = \int_0^1 (1-\lambda') \left(\frac{ds}{d\lambda'}\right)^2 d\lambda'. \qquad (14.61)$$

Аналогичным образом, меняя местами *P* и Q при неизменной соединительной линии (14.59), получаем

$$H^{Q/P} = \int_0^1 \lambda' \left(\frac{ds}{d\lambda'}\right)^2 d\lambda'. \qquad (14.62)$$

Сложим (14.62), (14.61) и применим неравенство Коши—Шварца:

$$\left[\int_0^1 \frac{ds}{d\lambda'} d\lambda'\right]^2 \leqslant \int_0^1 \left(\frac{ds}{d\lambda'}\right)^2 d\lambda'.$$

Это даст

$$\left[\int_0^1 \frac{ds}{d\lambda'} d\lambda'\right]^2 \leqslant H^{P/Q} + H^{Q/P},$$

а, следовательно, и (14.58), если учесть (14.55). Доказательство закончено.

14.4. Свойства энтропии в обобщенной версии

Определенная в предыдущем пункте энтропия (14.47), (14.50) обладает рядом свойств, аналогичных свойствам энтропии дискретной случайной величины, рассмотренным ранее. Такая аналогия является вполне естественной, если принять во внимание изложенную в п.14.3 интерпретацию энтропии (14.47) как асимптотического случая (при больших *N)* энтропии (14.34) дискретной случайной величины.

Свойство неотрицательности энтропии, для энтропии (14.47), (14.50) выполняется не всегда, а лишь при достаточно больших *N*. Условие

$$H_\xi^{P/Q} \leqslant \ln N$$

приводит к неотрицательности энтропии H_ξ.

В случае энтропии (14.47) при сравнении различных распределений *P* меру *ν* нужно оставлять фиксированной. Как указывалось, величина (14.51) является неотрицательной, поэтому из (14.60) имеем неравенство

$$H_\xi \leq \ln N.$$

В то же время, если положить P = Q, то, очевидно, будем иметь

$$H_\xi = \ln N.$$

Это доказывает следующее утверждение.

Теорема 10. *Энтропия* (14.47) *принимает максимальное значение, равное* ln *N, когда мера P пропорциональна мере* ν.

Этот результат является вполне естественным в свете данной в п. 14.3 дискретной интерпретации формулы (14.47). В самом деле, пропорциональность мер P и ν означает именно равномерное распределение вероятности по дискретным точкам A_i. Аналогом теоремы 10 для энтропии $H_\xi^{p/Q}$ является следующее утверждение: энтропия $H_\xi^{p/Q}$ принимает минимальное значение, равное нулю, когда распределение P совпадает с Q.

В соответствии с теоремой 10 энтропию $H_\xi^{p/Q}$, определенную формулой (14.51), целесообразно интерпретировать как дефект энтропии, т. е. как нехватку этой величины до ее максимального значения.

До сих пор предполагалось, что мера P является абсолютно непрерывной относительно меры Q или (что при конечных N то же самое) меры ν. Возникает вопрос, как определять энтропию H_ξ или $H_\xi^{p/Q}$, когда этой абсолютной непрерывности нет. Ответ на этот вопрос можно получить, рассматривая формулу (14.50) как асимптотический случай (при очень большом N) дискретной версии (14.34). Если, уплотняя точки A_i, о которых шла речь в п 14.3, вероятности $P(A_i)$ некоторых таких точек, вопреки формуле $P(A_i) \approx \Delta P/(N\Delta Q)$ (см. (14.38)), оставлять конечными: $P(A_i) > c > 0$ (c не зависит от N), то в пределе $N \to \infty$ мера Q_N не будет абсолютно непрерывна относительно меры P. Дефект $H_\xi^{p/Q}$ при $N \to \infty$ в этом случае будет неограниченно возрастать. Это дает основание считать, что

$$H_\xi^{P/Q} = \infty,$$

если мера P не является абсолютно непрерывной относительно Q (т. е. является сингулярной относительно ее). Сказанное, правда, еще не определяет энтропии H_ξ при отсутствии абсолютной непрерывности, так как для нее согласно (14.50) получается неопределенность типа $\infty - \infty$. Для ее определения требуется более детальный анализ предельного перехода $N \to \infty$, связанного с уплотнением точек A_i.

Прочие свойства энтропии в дискретной версии, касаются энтропии многих случайных величин и условной энтропии. При соответствующем определении последних понятий данные свойства будут иметь место и для обобщенной версии, основанной на определении (14.47).

Пусть имеются две случайные величины ξ, η. В соответствии с (14.47) они имеют энтропию

$$H_{\xi,\eta} = -\int \ln \frac{P(d\xi, d\eta)}{\nu(d\xi, d\eta)} P(d\xi, d\eta). \tag{14.63}$$

В то же время, применяя формулу (14.47) к одной случайной величине ξ или η, имеем

$$H_{\xi} = -\int \ln \frac{P(d\xi)}{\nu_1(d\xi)} P(d\xi)$$

$$H_{\eta} = -\int \ln \frac{P(d\eta)}{\nu_2(d\eta)} P(d\eta).$$

Здесь v_1, v_2 — некоторые меры; их связь с v рассматривается ниже. Условную энтропию $H_{\xi|\eta}$ определим как разность

$$H_{\xi|\eta} = H_{\xi\eta} - H_{\eta}, \tag{14.64}$$

т. е. так, чтобы было выполнено свойство аддитивности

$$H_{\xi\eta} = H_{\eta} + H_{\xi|\eta}. \tag{14.65}$$

Учитывая (14.47), (14.64), легко видеть, что для $H_{\xi|\eta}$ будем иметь формулу

$$H_{\xi|\eta} = -\int \ln \frac{P(d\xi \mid \eta)}{\nu(d\xi \mid \eta)} P(d\xi\, d\eta), \tag{14.66}$$

где $P\ (d\xi|\lambda\eta),\ \nu\ (d\xi|\eta)$ — условные меры, определяемые как производные Радона—Никодима при помощи обычных соотношений

$$P(\xi \in A, \eta \in B) = \int_{\eta \in B} P(\xi \in A \mid \eta) P(d\eta),$$

$$\nu(\xi \in A, \eta \in B) = \int_{\eta \in B} \nu(\xi \in A \mid \eta)\, \nu_2(d\eta)$$

(множества *A, B* произвольны). Определению (14.66) соответствует случайная энтропия

$$H(\xi \mid \eta) = -\ln \frac{P(d\xi \mid \eta)}{\nu(d\xi \mid \eta)}.$$

Данное выше определение условной энтропии $H_{\xi|\eta}$ можно многократно использовать при поэтапном рассмотрении цепочки случайных величин ξ_1, ξ_2, ..., ξ_n. При этом соотношение (14.65) приведет к выполнению свойства иерархической аддитивности:

$$H(\xi_1, \dots, \xi_n) = H(\xi_1) + H(\xi_2 \mid \xi_1) + H(\xi_3 \mid \xi_1, \xi_2) + \dots + H(\xi_n \mid \xi_1, \dots, \xi_{n-1}),$$

$$H_{\xi_1,\dots,\xi_n} = H_{\xi_1} + H_{\xi_2 | \xi_1} + H_{\xi_3 | \xi_1, \xi_2} + \dots + H_{\xi_n | \xi_1, \dots, \xi_{n-1}}, \tag{14.67}$$

14.5. Энтропия отрезка стационарного дискретного процесса и удельная энтропия

Будем предполагать, что имеется последовательность случайных величин ..., ξ_{k-1}, ξ_{k-}, ξ_{k+1}..... Индекс (параметр) k можно трактовать как дискретное время t, принимающее целочисленные значения ..., k— 1, k, k+ 1, ... Число различных значений индекса может быть неограниченно в обе стороны: $-\infty < k < \infty$; неограниченно в одну сторону, например: $0 < k < \infty$, или конечно: $1 \leq k \leq N$. Указанные значения образуют область *К,* определения параметра. Каждая случайная величина ξ_k пусть принимает одно из конечного или счетного числа значений, например, $\xi_k=(A_1, A_2, ..., A_m)$ (для дальнейшего существенна не столько конечность *m,* сколько конечность энтропии H_{ξ_k}). Описанный процесс — дискретную случайную величину как функцию дискретного параметра k, будем называть *дискретным процессом.*

Дискретный процесс является *стационарным,* если все законы распределения $P(\xi_{k_1}, ..., \xi_{k_r})$ (произвольной кратности r) не меняются при сдвиге:

$$\mathbf{P}[\xi_{k_1} = x_1, \dots, \xi_{k_r} = x_r] = \mathbf{P}[\xi_{k_1+a} = x_1, \dots, \xi_{k_r+a} = x_r] \tag{14.68}$$

(a — любое целое).

Величина сдига a предполагается такой, что значения $k_1+a, ..., k_r+a$ не выходят из области K определения параметра k. В дальнейшем мы это не будем оговаривать, предполагая, скажем, что область определения параметра не ограничена в обе стороны.

Рассмотрим различные условные энтропии одной из случайных величин ξ_k дискретного стационарного случайного процесса. Ее безусловная энтропия

$$H_{\xi_k} = -\sum_{\xi_k} P(\xi_k) \ln P(\xi_k)$$

вследствие свойства (14.68) не зависит от выбранного значения параметра k.

Аналогично согласно (14.68) условная энтропия $H_{\xi_k|\xi_{k-1}}$ не зависит от k. Применяя теорему о том, что условная энтропия не может превосходить безусловную, имеем неравенство

$$H_{\xi_k|\xi_{k-1}} \leqslant H_{\xi_k}$$

или, учитывая стационарность,

$$H_{\xi_2|\xi_1} = H_{\xi_3|\xi_2} = \dots = H_{\xi_k|\xi_{k-1}} \leqslant H_{\xi_1} = H_{\xi_2} = \dots = H_{\xi_k}.$$

Если ввести условную энтропию $H_{\xi_k \mid \xi_{k-1}\,\xi_{k-2}}$, то применение теоремы о том, что при добавлении условий условная энтропия не увеличивается, при $\xi = \xi_k$, $\eta = \xi_{k-1}$, $\zeta = \xi_{k-2}$ даст неравенство

$$H_{\xi_k \mid \xi_{k-1},\, \xi_{k-2}} \leqslant H_{\xi_k \mid \xi_{k-1}}.$$

Согласно условию стационарности (14.68) эта энтропия не зависит от k.

Аналогичным образом, увеличивая число случайных величин в условии, мы и дальше будем иметь монотонное изменение (невозрастание) условной энтропии:

$$H_{\xi_k} \geqslant H_{\xi_k \mid \xi_{k-1}} \geqslant H_{\xi_k \mid \xi_{k-1},\, \xi_{k-2}} \geqslant \ldots \geqslant$$
$$\geqslant H_{\xi_k \mid \xi_{k-1}, \ldots, \xi_{k-l}} \geqslant \ldots \geqslant 0. \tag{14.69}$$

Кроме того, все условные энтропии неотрицательны, т. е. ограничены снизу. Отсюда вытекает существование неотрицательного предела

$$H_1 = \lim_{l \to \infty} H_{\xi_k \mid \xi_{k-1}, \ldots, \xi_{k-l}}, \tag{14.70}$$

который мы будем обозначать также h, чтобы избежать увеличения числа индексов.

Этот предел мы определяем как *удельную энтропию,* рассчитанную на один элемент последовательности $\{\xi_k\}$. Основанием для такого наименования является следующая теорема.

Теорема 11. *Если* $\{\xi_k\}$ *— стационарный дискретный процесс, такой, что* $H_{\xi_k} < \infty$ *то предел*

$$\lim_{l \to \infty} H_{\xi_1 \ldots \xi_l} / l$$

существует и равен пределу (14.70).

Доказательство. Рассмотрим энтропию $H_{\xi_1 \ldots \xi_{m+n}}$ и представим ее в виде

$$H_{\xi_{m+n} \ldots \xi_1} = H_{\xi_m \ldots \xi_1} + H_{\xi_{m+1} \mid \xi_m \ldots \xi_1} + H_{\xi_{m+2} \mid \xi_{m+1} \ldots \xi_1} +$$
$$+ \ldots + H_{\xi_{m+n} \mid \xi_{m+n-1} \ldots \xi_1}. \tag{14.71}$$

Поскольку

$$H_{\xi_{m+i} \mid \xi_{m+i-1} \ldots \xi_1} = H_1 + o_{m+i}(1) = H_1 + o_m(1) \qquad (i \geqslant 1)$$

согласно (14.70) (здесь $o_j(1) \to 0$ при $j \to \infty$), то из (14.71) имеем (после деления на $m + n$)

$$\frac{H_{\xi_{m+n}\dots\xi_1}}{m+n} = \frac{H_{\xi_m\dots\xi_1}}{m+n} + \frac{n}{m+n}H_1 + \frac{n}{m+n}o_m(1). \tag{14.72}$$

Пусть m и n стремятся к бесконечности так, что $n/m \to\infty$; тогда $n/(m+n)$ будет стремиться к 1, а отношение $H_{\xi_m\dots\xi_1}/(m+n)$, которое можно оценить

$$\frac{1}{m+n}H_{\xi_m\dots\xi_1} \leqslant \frac{m}{m+n}H_{\xi_1},$$

будет, очевидно, стремиться к 0.

Поэтому из равенства (14.72) получаем утверждение теоремы. Доказательство закончено.

Нетрудно доказать также, что с ростом l отношение $H_{\xi_1\dots\xi_l}/l$ меняется монотонно, т. е. не возрастает. Для этого образуем разность

$$\delta = \frac{1}{l}H_{\xi_1\dots\xi_l} - \frac{1}{l+1}H_{\xi_1\dots\xi_{l+1}} = \frac{1}{l}H_{\xi_1\dots\xi_l} - \frac{1}{l+1}\left[H_{\xi_1\dots\xi_l} + H_{\xi_{l+1}|\xi_1\dots\xi_l}\right] = \frac{1}{l(l+1)}H_{\xi_1\dots\xi_l} - \frac{1}{l+1}H_{\xi_{l+1}|\xi_1\dots\xi_l}$$

и представим ее в виде

$$\delta = \frac{1}{l(l+1)}\left[H_{\xi_1\dots\xi_l} - lH_{\xi_{l+1}|\xi_1\dots\xi_l}\right] = \frac{1}{l(l+1)}\sum_{i=1}^{l}\left[H_{\xi_i|\xi_1\dots\xi_{i-1}} - H_{\xi_i|\xi_{i-l}\dots\xi_{i-1}}\right]. \tag{14.73}$$

Вследствие неравенств (14.69) слагаемые в правой части (14.73) являются неотрицательными, откуда вытекает неотрицательность разности δ.

Согласно теореме 11 имеет место равенство

$$H_{\xi_1\dots\xi_n} = nH_1 + no_n(1). \tag{14.74}$$

Образуем комбинацию

$$H_{\xi_1\dots\xi_m} + H_{\xi_1\dots\xi_n} - H_{\xi_1\dots\xi_{m+n}} = H_{\xi_1\dots\xi_m} + H_{\xi_{m+1}\dots\xi_{m+n}} - H_{\xi_1\dots\xi_{m+n}}. \tag{14.75}$$

Легко доказать, что она является монотонной функцией от m при фиксированном n (или наоборот). В самом деле, ее можно записать в форме

$$H_{\xi_{m+1}\dots\xi_{m+n}}-H_{\xi_{m+1}\dots\xi_{m+n}\mid\xi_1\dots\xi_m}. \tag{14.76}$$

Поскольку условные энтропии не возрастают с ростом m согласно (14.69), то выражение (14.76) не убывает при увеличении m. То же самое можно сказать и о зависимости от n, так как в (14.75) m и n входят симметрично. Из (5.1.9) очевидна также неотрицательность комбинации (5.1.8)

Рассмотрим предел

$$2\Gamma=\lim_{n\to\infty}\lim_{m\to\infty}\left[H_{\xi_1\dots\xi_m}+H_{\xi_1\dots\xi_n}-H_{\xi_1\dots\xi_{m+n}}\right]=$$
$$=\lim_{m\to\infty}\lim_{n\to\infty}\left[H_{\xi_1\dots\xi_m}+H_{\xi_1\dots\xi_n}-H_{\xi_1\dots\xi_{m+n}}\right]. \tag{14.77}$$

Менять порядок пределов здесь можно вследствие уже отмеченной симметрии относительно перестановки m и n. В силу указанной монотонности этот предел (конечный или бесконечный) всегда существует. Переходя от формы записи (14.76) к форме (14.76), с использованием иерархического соотношения

$$H_{\xi_{m+1}\dots\xi_{m+n}\mid\xi_1\dots\xi_m}=\sum_{i=1}^{n}H_{\xi_{m+i}\mid\xi_1\dots\xi_{m+i-1}}$$

равенство (14.77) можно записать, выполнив предельный переход $m\to\infty$:

$$2\Gamma=\lim_{n\to\infty}\left[H_{\xi_1\dots\xi_n}-nH_1\right], \tag{14.78}$$

так как

$$H_{\xi_{m+i}\mid\xi_1\dots\xi_{m+i-1}}\to H_1 \text{ при } m\to\infty.$$

Если энтропию $H_{\xi_1\dots\xi_n}$ здесь представить в виде иерархической суммы, то формула (14.78) примет вид

$$2\Gamma=\sum_{j=0}^{\infty}\left[H_{\xi_{j+1}\mid\xi_1\dots\xi_j}-H_1\right]. \tag{14.79}$$

Согласно (14.78) имеем

$$H_{\xi_1\dots\xi_n}=nH_1+2\Gamma+o_n(1). \tag{14.80}$$

Эта формула уточняет (14.74).

Поскольку на один элемент последовательности $\{\xi_k\}$ приходится в среднем энтропия H_1, то на n таких элементов приходится nH_1. Согласно (14.80) энтропия 2Г отличается от nH_1 при больших n на величину 2Г, которую можно трактовать как энтропию, приходящуюся на концы отрезка. Таким образом, *энтропия каждого конца отрезка стационарного процесса в пределе равна* Г.

Если один длинный отрезок стационарного процесса разбить на два отрезка и ликвидировать статистические связи (корреляции) между процессами на этих отрезках, то энтропия возрастет приблизительно на 2Г, так как появятся два новых конца.

Применяя формулу (14.80) для трех отрезков длиной *m, n* и *m* + *n* и образуя комбинацию (14.75), теперь имеем

$$H_{\xi_1 \dots \xi_m} + H_{\xi_1 \dots \xi_n} - H_{\xi_1 \dots \xi_{m+n}} = 2\Gamma + o_m(1) + o_n(1) + o_{m+n} \quad (*)$$

Это согласуется с (14.77) и подтверждает сказанное об увеличении энтропии на 2Г.

14.6. Энтропия марковской цепи

1. Пусть дискретный (необязательно стационарный) процесс $\{\xi_k\}$ является марковским. Это значит, что совместные законы распределения последовательно идущих случайных величин распадаются на произведение

$$P(\xi_k, \xi_{k+1}, \dots, \xi_{k+l}) = \\ = P(\xi_k)\, \pi_k(\xi_k, \xi_{k+1}) \dots \pi_{k+l-1}(\xi_{k+l-1}, \xi_{k+l}) \quad (14.81)$$

функций $\pi_j(\xi, \xi') = P(\xi' \mid \xi)$, которые называются вероятностями перехода. В (14.81) вероятности $P(\xi_k)$ соответствуют однократному распределению случайной величины ξ_k.

Вероятность перехода $\pi_j(\xi_j, \xi_{j+1})$, равняясь условным вероятностям $P(\xi_{j+1} \mid \xi_j)$, обладает свойством неотрицательности и нормировки

$$\sum_{\xi'} \pi_j(\xi, \xi') = 1. \quad (14.82)$$

Дискретный марковский процесс называют также марковской цепью.

Переходя от вероятностей (14.81) к условным вероятностям, легко получить, что в случае марковского процесса

$$P(\xi_{k+m+1} \mid \xi_k, \dots, \xi_{k+m}) = \pi_{k+m}(\xi_{k+m}, \xi_{k+m+1}) \quad (\text{любое } m \geqslant 1)$$

и, следовательно,

$$P(\xi_{k+m+1} \mid \xi_k, \dots, \xi_{k+m}) = P(\xi_{k+m+1} \mid \xi_{k+m}).$$

Поэтому

$$H(\xi_{k+m+1} \mid \xi_k, \dots, \xi_{k+m}) = -\ln \pi_{k+m}(\xi_{k+m}, \xi_{k+m+1}) = \\ = H(\xi_{k+m+1} \mid \xi_{k+m}). \quad (14.83)$$

Благодаря (14.83) иерархическая формула принимает вид

$$H(\xi_1,\ldots,\xi_n)=H(\xi_1)+H(\xi_2\,|\,\xi_1)+H(\xi_3\,|\,\xi_2)+\ldots+H(\xi_n\,|\,\xi_{n-1}). \tag{14.84}$$

Аналогичным образом для средних энтропии имеем

$$H_{\xi_{k+m+1}\,|\,\xi_k\ldots\xi_{k+m}}=H_{\xi_{k+m+1}\,|\,\xi_{k+m}}, \tag{14.85}$$

$$H_{\xi_1\ldots\xi_n}=H_{\xi_1}+H_{\xi_2\,|\,\xi_1}+H_{\xi_3\,|\,\xi_2}+\ldots+H_{\xi_n\,|\,\xi_{n-1}}. \tag{14.86}$$

Дискретный марковский процесс является стационарным, если вероятности перехода π_k (ξ, ξ') не зависят от значения параметра k и если все однократные распределения $P(\xi_k)$ одинаковы (равны $P_{\text{ст}}(\xi)$). Образуя многократные распределения по формуле (14.81), легко доказать, что они также удовлетворяют условию стационарности (14.68).

Стационарность однократного распределения $P_{\text{ст}}(\xi)$ означает, что выполняется уравнение

$$\sum_{\xi}P_{\text{ст}}(\xi)\,\pi(\xi,\,\xi')=P_{\text{ст}}(\xi'). \tag{14.87}$$

Последнее легко вывести, записывая согласно (14.81) двукратное распределение

$$P(\xi_k,\,\xi_{k+1})=P_{\text{ст}}(\xi_k)\,\pi(\xi_k,\,\xi_{k+1})$$

и суммируя его по $\xi_k=\xi$, что дает $P(\xi_{k+1})$.

Учитывая (14.83), легко видеть, что для стационарного марковского процесса удельная энтропия (14.70) совпадает с энтропией, соответствующей вероятностям перехода $\pi(\xi,\,\xi')$ при стационарном распределении вероятностей:

$$H_1=-\sum_{\xi}P_{\text{ст}}(\xi)\sum_{\xi'}\pi(\xi,\,\xi')\ln\pi(\xi,\,\xi'). \tag{14.88}$$

В самом деле, усредняя (14.83) со стационарными вероятностями $P_{\text{ст}}(\xi)\ \pi(\xi,\,\xi')$, убеждаемся, что всем значениям $m\geq 1$ соответствует одна и та же условная энтропия $H_{\xi_{k+m+1}\,|\,\xi_{k+1}\ldots\xi_{k+m}}$, равная (14.88).

Усредним далее формулу (14.84) со стационарными вероятностями. Это даст равенство

$$H_{\xi_1\ldots\xi_n}=H_{\xi_1}+(n-1)H_1. \tag{14.89}$$

Сравнивая его с (14.80), видим, что в стационарном марковском случае

$$2\Gamma=H_{\xi_1}-H_1=\sum_{\xi,\,\xi'}P_{\text{ст}}(\xi)\,\pi(\xi,\,\xi')\ln\frac{\pi(\xi,\,\xi')}{P_{\text{ст}}(\xi')} \tag{14.90}$$

и $o_n(1)=0$, т. е. формула

$$H_{\xi_1 \ldots \xi_n} = nH_1 + 2\Gamma \tag{14.91}$$

выполняется точно. К результату (14.90) можно прийти также при помощи формулы (14.79). В самом деле, поскольку, как отмечалось,

$$H_{\xi_{j+1} \mid \xi_1 \ldots \xi_j} = H_1,$$

то в сумме в правой части (*) остается лишь единственный ненулевой член:

$$2\Gamma = H_{\xi_1} - H_1,$$

что совпадает с (14.90).

2. Итак, для вычисления энтропии, если задана матрица вероятностей перехода

$$\pi = \| \pi(\xi, \xi') \|, \tag{14.92}$$

следует найти стационарное распределение и воспользоваться формулами (14.88), (14.89). Уравнение (14.87) вполне однозначно определяет стационарное распределение $P_{ст}(\xi)$, если марковский процесс является эргодическим, т. е. если собственное значение $\lambda = 1$ является невырожденным собственным значением матрицы (14.92). Согласно теореме о разложении определителей, из уравнения $\det(\pi—1) = 0$ будет вытекать (14.87), если положить

$$P_{ст}(\xi) = A_{\xi\xi'} \Big/ \sum_{\xi''} A_{\xi''\xi'}, \tag{14.93}$$

где $A_{\xi\xi'}$ — алгебраические дополнения элемента $a_{\xi\xi'}$ матрицы

$$\| a_{\xi\xi'} \| = \| \pi(\xi, \xi') - \delta_{\xi\xi'} \| \equiv \pi - 1.$$

Значение ξ' здесь любое, так что достаточно вычислить алгебраические дополнения одного столбца указанной матрицы.

Если марковский процесс неэргодический, то стационарное распределение не определяется полностью матрицей перехода, а зависит еще от начального (или какого-либо другого однократного) распределения. В этом случае в формуле (14.93) алгебраические дополнения равны нулю и, следовательно, имеет место неопределенность типа 0/0. Стационарное распределение при этом может быть выражено через миноры меньшего порядка и начальное распределение.

Как известно, состояния дискретного марковского процесса можно расположить в таком порядке, чтобы переходная матрица имела следующий «ящичный» вид:

$$\pi = \begin{pmatrix} \pi^{00} & \pi^{01} & \pi^{02} & \pi^{03} & \dots \\ 0 & \pi^{11} & 0 & 0 & \dots \\ 0 & 0 & \pi^{22} & 0 & \dots \\ & & & & \end{pmatrix} = \Pi^{00} + \sum_{\iota} (\Pi^{0\iota} + \Pi^{\iota\iota}), \qquad (14.94)$$

где

$$\Pi^{00} = \begin{pmatrix} \pi^{00} & 0 & \dots \\ 0 & 0 & \dots \\ & \vdots & \end{pmatrix}, \qquad \Pi^{01} = \begin{pmatrix} 0 & \pi^{01} & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix},$$

и т. д.

Здесь π^{ij} означает матрицу, имеющую размерность $r_i \times r_j$ и описывающую переход из подмножества E_i, содержащего r_i состояний, в подмножество E_j с r_j состояниями. На прочих местах стоят нули. Множества E_1, E_2, ... образуют эргодические классы. Из множества E_0 происходят переходы в эргодические классы E_1, E_2, ... Из каждого такого класса нет выхода. Поэтому внутри каждого класса устанавливается свое стационарное распределение. Оно может быть найдено по формуле типа (14.93)

$$P^{\iota}_{\text{ст}}(\xi) = A^{\iota\iota}_{\xi\xi'} \Big/ \sum_{\xi'' \in E_{\iota}} A^{\iota\iota}_{\xi''\xi'} \quad \text{при } \xi \in E_{\iota} \qquad (14.95)$$

с той лишь разницей, что теперь берутся алгебраические дополнения подматрицы $\| \pi^{\iota\iota}_{\xi\xi'} - \delta_{\xi\xi'} \|, \ \xi, \ \xi' \in E_{\iota}$, которые уже не равны нулю. Вероятности $P'_{CT}(\xi)$ сосредоточены на E_1. Полное стационарное распределение есть линейная комбинация

$$P_{\text{ст}}(\xi) = \sum q_i P^{i}_{\text{ст}}(\xi) \qquad (14.96)$$

частных распределений (14.95), которые, как легко видеть, являются ортогональными

$$\sum_{\xi} P^{i}_{\text{ст}}(\xi) P^{j}_{\text{ст}}(\xi) = 0 \qquad \text{при } i \neq j.$$

Коэффициенты q_i этой линейной комбинации определяются начальным распределением $P(\xi_1) = P_1(\xi_1)$. Они совпадают с результирующей вероятностью попадания в тот или иной эргодический класс и удовлетворяют условию нормировки

$$\sum_i q_i = 1.$$

Принимая во внимание вид (14.94) переходной матрицы и суммируя переходы из E_0 в E_i на различных этапах, можно найти, как они явно выражаются через начальное распределение:

$$q_i = \sum_{\xi \in E_i} P_1(\xi) + \sum_{\xi\xi'} P_1(\xi)\,[\Pi^{0i}_{\xi\xi'} +$$
$$+ (\Pi^{00}\,\Pi^{0i})_{\xi\xi'} + (\Pi^{00}\,\Pi^{00}\,\Pi^{0i})_{\xi\xi'} + \ldots],\ i \neq 0.$$

В круглых скобках здесь стоит матричное произведение. Суммируя степени матрицы Π^{00}, имеем

$$q_i = \sum_{\xi \in E_i} P_1(\xi) + \sum_{\xi,\ \xi'} P_1(\xi)\,([1 - \Pi^{00}]^{-1}\,\Pi^{0i})_{\xi\xi'},\ q_0 = 0.$$

Учитывая (14.96), (14.94), получаем, что удельную энтропию (14.88) удобно представить суммой

$$H_1 = \sum q_i H_1^i, \tag{14.97}$$

где частные энтропии

$$H_1^i = - \sum_{\xi,\ \xi' \in E_i} P^i_{\text{ст}}(\xi)\,\pi^{ii}(\xi,\ \xi')\ln \pi^{ii}(\xi,\ \xi') \tag{14.98}$$

вычисляются совершенно аналогично эргодическому случаю. Причина этого в том, что неэргодический процесс является статистической смесью (с вероятностями q_i) эргодических процессов, имеющих меньшее число состояний r_i.

Суммирование в (14.97), (14.98) проводится только по эргодическим классам E_1, E_2.... . Подпространство же E_0 имеет нулевую стационарную вероятность. Соединение всех эргодических классов $E_1+E_2+ \ldots = E_a$, на котором концентрируется стационарная вероятность, можно назвать «активным» подпространством. На удельную энтропию H_1 распределения и переходы в «пассивном» пространстве E_0 оказывают влияние лишь постольку, поскольку они влияют на вероятности q_i. Если процесс эргодический, т. е. имеется, кроме E_0, лишь один эргодический класс E_1, тогда q_1=1 и пассивное пространство вообще не оказывает никакого влияния на удельную энтропию. В этом случае удельная энтропия марковского процесса в пространстве $E_0 + E_a$ совпадает с удельной энтропией марковского процесса, протекающего в подпространстве $E_a = E_1$ и имеющего матрицу перехода π^{11}.

3. Чтобы проиллюстрировать применение формул, полученных в этом пункте, рассмотрим несколько несложных примеров.

Пример 1. Рассмотрим сначала простейший дискретный марковский процесс — процесс с двумя состояниями. Матрица (14.92) является при этом 2×2-матрицей. Вследствие условия нормировки (14.82) ее элементы не являются независимыми. Имеются лишь два независимых параметра μ и ν, определяющих матрицу π:

$$\pi = \begin{pmatrix} 1-\mu & \mu \\ \nu & 1-\nu \end{pmatrix}.$$

Поскольку в данном случае

$$a = \pi - 1 = \begin{pmatrix} -\mu & \mu \\ \nu & -\nu \end{pmatrix}; \quad A_{11} = -\nu; \quad A_{21} = -\mu,$$

то согласно (14.93) имеем стационарное распределение

$$P_{\text{ст}}(1) = \frac{\nu}{\mu+\nu}; \qquad P_{\text{ст}}(2) = \frac{\mu}{\mu+\nu} \tag{14.99}$$

и кроме того

$$\| P_{\text{ст}}(\xi_k, \xi_{k+1}) \| = \frac{1}{\mu+\nu} \begin{pmatrix} \nu-\mu\nu & \mu\nu \\ \mu\nu & \mu-\mu\nu \end{pmatrix}.$$

Применяя формулу (14.88), находим удельную энтропию

$$H_1 = \frac{\nu}{\mu+\nu} h_2(\mu) + \frac{\mu}{\mu+\nu} h_2(\nu), \tag{14.100}$$

где

$$h_2(x) = -x \ln x - (1-x) \ln (1-x). \tag{14.101}$$

Далее по формуле (14.100) нетрудно получить граничную энтропию

$$2\Gamma = h_2\left(\frac{\mu}{\mu+\nu}\right) - \frac{\nu h_2(\mu) + \mu h_2(\nu)}{\mu+\nu}.$$

Пример 2. Пусть теперь имеется процесс с тремя состояниями, имеющий переходную матрицу

$$\pi = \begin{pmatrix} 1-\mu & \mu' & \mu'' \\ \nu'' & 1-\nu & \nu' \\ \lambda' & \lambda'' & 1-\lambda \end{pmatrix} \qquad (\mu = \mu' + \mu'' \text{ и т. д.})$$

При этом, очевидно,

$$a = \begin{pmatrix} -\mu & \mu' & \mu'' \\ \nu'' & -\nu & \nu' \\ \lambda' & \lambda'' & -\lambda \end{pmatrix}.$$

По формуле (14.93) находим стационарное распределение

$$P_{\text{ст}}(\xi) = A_{\xi 1} / (A_{11} + A_{21} + A_{31}), \tag{14.102}$$

где

$$A_{11}=\begin{vmatrix} -\nu & \nu' \\ \lambda'' & -\lambda \end{vmatrix}=\lambda\nu-\lambda''\nu'=\lambda'\nu'+\lambda'\nu''+\lambda''\nu'';$$

$$A_{21}=\mu'\lambda'+\mu'\lambda''+\lambda''\mu'';$$

$$A_{31}=\mu'\nu'+\nu'\mu''+\nu''\mu''.$$

Удельная энтропия (14.88) оказывается равной

$$H_1=P_{\text{ст}}(1)h_3(\mu',\ \mu'')+P_{\text{ст}}(2)h_3(\nu',\ \nu'')+P_{\text{ст}}(3)h_3(\lambda',\ \lambda''),$$

где обозначено

$$h_3(\mu',\ \mu'')=-\mu'\ln\mu'-\mu''\ln\mu''-(1-\mu'-\mu'')\ln(1-\mu'-\mu''). \qquad (14.103)$$

Данный процесс с тремя состояниями будет неэргодическим, если, например, $\lambda'=\lambda''=0$; $\mu''=\nu'=0$, так что матрица перехода имеет «ящичный» вид

$$\pi=\begin{pmatrix} 1-\mu & \mu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

При такой матрице третье состояние остается неизменным, а переходы совершаются только между первым и вторым состояниями. Алгебраические дополнения (14.102) обращаются в нуль. Стационарными, как легко понять, являются следующие распределения:

$$P^1_{\text{ст}}(\xi)=\begin{cases} \nu/(\mu+\nu) & \text{при } \xi=1; \\ \mu/(\mu+\nu) & \text{при } \xi=2; \\ 0 & \text{при } \xi=3; \end{cases} \qquad P^2_{\text{ст}}(\xi)=\delta_{\xi 3}.$$

Первое из них совпадает с (14.99). Указанные функции $P^1_{\text{ст}}(\xi)$, $P^2_{\text{ст}}(\xi)$ являются ортогональными. Задаваясь начальным распределением $P_1(\xi)$, по формуле (14.96) находим результирующее стационарное распределение

$$P_{\text{ст}}(\xi)=[P_1(1)+P_1(2)]\,P^1_{\text{ст}}(\xi)+P_1(3)\,P^2_{\text{ст}}(\xi).$$

Теперь согласно (14.97) нетрудно записать удельную энтропию

$$H_1=[P_1(1)+P_1(2)]\,(\nu h(\mu)+\mu h(\nu))/(\mu+\nu).$$

14.7. Удельная энтропия части компонент дискретного марковского процесса и условного марковского процесса

1. Вычислим удельную энтропию стационарного немарковского дискретного процесса $\{y_k\}$, который, однако, может быть дополнен до марковского процесса присоединением добавочных компонент x_k. Совокупность $\xi=(x, y)$ будет образовывать фазовое пространство марковского процесса, и процесс $\{\xi_k\} = \{x_k, y_k)$ уже будет марковским. К нему можно применять результаты предыдущего пункта и, в частности, находить удельную энтропию, которую будем обозначать $H^1_{xy}=h_{xy}$. Энтропия $H_{y_1 \dots y_n}$ исходного процесса отличается от энтропии

$$H_{\xi_1 \dots \xi_n} = H_{y_1 \dots y_n} + H_{x_1 \dots x_n | y_1 \dots y_n} \qquad (14.104)$$

марковского процесса на условную энтропию $H_{x_1 \dots x_n | y_1 \dots y_n}$ называемую *энтропией условного марковского процесса* $\{x_k\}$ (при, фиксированном $\{y_k\}$).

Наряду с удельной энтропией $h_{xy} = h_\xi$ стационарного марковского процесса введем удельную энтропию исходного y-процесса

$$h_y = \lim_{n \to \infty} \frac{1}{n} H_{y_1 \dots y_n} \qquad (14.105)$$

и условного x-процесса

$$h_{x|y} = \lim_{n \to \infty} \frac{1}{n} H_{x_1 \dots x_n | y_1 \dots y_n}. \qquad (14.106)$$

Они вследствие (14.104) связаны с h_{xy} соотношением

$$h_y + h_{x|y} = h_{xy}. \qquad (14.107)$$

Нетрудно понять, что можно также рассматривать x-процесс как априорный немарковский (в общем случае) процесс и условный y-процесс при фиксированном x. Им будут соответствовать удельные энтропии

$$h_x = \lim_{n \to \infty} \frac{1}{n} H_{x_1 \dots x_n}; \qquad h_{y|x} = \lim_{n \to \infty} H_{y_1 \dots y_n | x_1 \dots x_n}. \qquad (14.108)$$

Соотношения (14.104), (14.107) при этом заменяются на

$$H_{x_1 \dots x_n} + H_{y_1 \dots y_n | x_1 \dots x_n} = H_{\xi_1 \dots \xi_n}; \qquad (14.109)$$

$$h_x + h_{y|x} = h_{xy}. \qquad (14.110)$$

Поскольку энтропию марковского процесса мы уже умеем находить, достаточно научиться вычислять лишь одну из величин h_y или $h_{x|y}$.

Вторая находится при помощи (14.107). Вследствие симметрии соответствующая величина из h_x, $h_{y|x}$ будет вычисляться тем же способом, а вторая — из соотношения (14.110).

Излагаемый ниже (п.2) метод можно применять для вычисления энтропии условного марковского процесса как в стационарном, так и в нестационарном случае. Стационарные же распределения вероятностей и пределы (14.105), (14.106), (14.108), как правило, существуют только в стационарном случае.

Условную энтропию $H_{x_1 \dots x_n | y_1 \dots y_n}$ можно представить в виде суммы

$$H_{x_1 \dots x_n | y_1 \dots y_n} = H_{x_1 | y_1 \dots y_n} + H_{x_2 | x_1 y_1 \dots y_n} + \\ + H_{x_3 | x_1 x_2 y_1 \dots y_n} + \dots + H_{x_n | x_1 \dots x_{n-1} y_1 \dots y_n}$$

и предел (14.106) заменить на предел

$$h_{x|y} = \lim_{k \to \infty,\ n-k \to \infty} H_{x_k | x_1 \dots x_{k-1}, y_1 \dots y_n} = \\ = H_{x_k | \dots \xi_{k-2} \xi_{k-1} y_k y_{k+1} \dots} \tag{14.111}$$

Об эквивалентности двух таких представлений удельной энтропии была речь в теореме 11. Конечно, теперь процесс $\{x_k\}$ является условным и поэтому нестационарным. Прямо применять к нему теорему 11. нельзя, поэтому необходимо произвести ее обобщение.

Теорема 12. *Для стационарного процесса* $\{\xi_k\}=\{x_k, y_k\}$ *пределы (14.106) и (14.111) совпадают.*

Доказательство. Поскольку

$$\lim_{n-k \to \infty} H_{x_k | x_1 \dots x_{k-1} y_1 \dots y_n} = H_{x_k | \xi_1 \dots \xi_{k-1} y_k y_{k+1} \dots},$$

$$\lim_{k \to \infty} H_{x_k | \xi_1 \dots \xi_{k-1} y_k y_{k+1} \dots} = H_{x_k | \dots \xi_{k-2} \xi_{k-1} y_k y_{k+1} \dots},$$

то имеем

$$H_{x_k | x_1 \dots x_{k-1} y_1 \dots y_n} = H_{x_k | \dots \xi_{k-2} \xi_{k-1} y_k y_{k+1} \dots} + \\ + o_k(1) + o_{n-k}(1) \tag{14.112}$$

(предполагается, что оценка $o_{n-k}(1)$ равномерна по k).

Представим n в виде суммы трех чисел: $n = m + r + s$. Подставляя (14.112) в равенство

$$H_{x_1 \dots x_n | y_1 \dots y_n} = H_{x_1 \dots x_m | y_1 \dots y_n} + H_{x_{m+1} | \xi_1 \dots \xi_m y_{m+1} \dots y_n} + \\ + \dots + H_{x_{m+r} | \xi_1 \dots \xi_{m+r-1} y_{m+r} \dots y_n} + \\ + H_{x_{m+r+1} \dots x_n | \xi_1 \dots \xi_{m+r} y_{m+r+1} \dots y_n},$$

получаем

$$\frac{1}{n} H_{x_1 \ldots x_n | y_1 \ldots y_n} = \frac{1}{m+r+s} H_{x_1 \ldots x_m | y_1 \ldots y_n} +$$
$$+ \frac{r}{m+r+s} H_{x_k | \ldots \xi_{k-1} y_k \ldots} + \frac{r}{m+r+s} [o_m(1) + o_s(1)] +$$
$$+ \frac{1}{m+r+s} H_{x_{m+r+1} \ldots x_n | \xi_1 \ldots \xi_{m+r} y_{m+r+1} \ldots y_n}. \qquad (14.113)$$

Здесь

$$H_{x_1 \ldots x_m | y_1 \ldots y_n} = \sum_{k=1}^{m} H_{x_k | x_1 \ldots x_{k-1} y_1 \ldots y_n} \leqslant m H_{x_k}$$

и

$$H_{x_{m+r+1} \ldots x_n | \xi_1 \ldots \xi_{m+r} y_{m+r+1} \ldots y_n} \leqslant s H_{x_k},$$

так как условная энтропия не больше безусловной. Поэтому, если в (14.113) совершить предельный переход $m \to \infty$, $r \to \infty$, $s \to \infty$, причем $r/m \to \infty$; $r/s \to \infty$, то в этом выражении останется лишь один член $H_{x_k | \ldots, \xi_{k-1}, y_k \ldots}$. Это доказывает теорему.

Как видно из приведенного доказательства, теорема 12 справедлива не только в случае марковского совокупного процесса $\{x_k, y_k\}$. Для марковского же процесса вследствие условия Маркова, имеем

$$P(x_k | x_1, \ldots, x_{k-1} y_1, \ldots, y_n) =$$
$$= P(x_k | x_{k-1}, y_{k-1}, y_k, \ldots, y_n) \quad (n \geqslant k)$$

и

$$H_{x_k | x_1 \ldots x_{k-1}, y_1 \ldots y_n} = H_{x_k | \xi_{k-1} y_k \ldots y_n}.$$

Следовательно, формула (14.111) принимает вид

$$h_{x|y} = H_{x_k | x_{k-1} y_{k-1} y_k y_{k+1} \ldots}.$$

2. Перейдем к вычислению энтропии y-процесса и условного процесса для марковского совокупного процесса. Для этого рассмотрим условную энтропию

$$H_{y_k | y_1 \ldots y_{k-1}} = -\mathbf{M} \ln P(y_k | y_1, \ldots, y_{k-1}), \qquad (14.114)$$

определяющую согласно (14.70) в пределе

$$h_y = \lim_{k \to \infty} H_{y_k | y_1 \ldots y_{k-1}} \qquad (14.115)$$

удельную энтропию h_y.

Пользуясь условием марковости (14.81), запишем многомерное распределение вероятностей

$$P(\xi_1, \ldots, \xi_n) = P(\xi_1) \pi(\xi_1, \xi_2) \ldots \pi(\xi_{k-1}, \xi_k),$$

где ξ_j обозначает пару x_j, y_j.

Применяя формулу обратной вероятности (формулу Бейеса), отсюда получаем

$$P(y_k \mid y_1, \ldots, y_{k-1}) = \frac{\sum\limits_{x_1 \ldots x_k} P(\xi_1)\pi(\xi_1, \xi_2) \ldots \pi(\xi_{k-1}, \xi_k)}{\sum\limits_{x_1 \ldots x_{k-1}} P(\xi_1)\pi(\xi_1, \xi_2) \ldots \pi(\xi_{k-2}, \xi_{k-1})}. \qquad (14.116)$$

Следуя теории условных марковских процессов, введем *финальные апостериорные вероятности*

$$P(x_{k-1} \mid y_1, \ldots, y_{k-1}) \equiv W_{k-1}(x_{k-1}) = \frac{\sum\limits_{x_1 \ldots x_{k-2}} P(\xi_1)\pi(\xi_1, \xi_2) \ldots \pi(\xi_{k-2}, \xi_{k-1})}{\sum\limits_{x_1 \ldots x_{k-1}} P(\xi_1)\pi(\xi_1, \xi_2) \ldots \pi(\xi_{k-2}, \xi_{k-1})}. \qquad (14.117)$$

При помощи них выражение (14.116) запишется так:

$$P(y_k \mid y_1, \ldots, y_{k-1}) = \sum_{x_{k-1}, x_k} W_{k-1}(x_{k-1})\pi(x_{k-1}, y_{k-1}; x_k, y_k)$$

или

$$P(y_k \mid y_1, \ldots, y_{k-1}) = \sum_x W_{k-1}(x)\pi(x, y_{k-1}; y_k), \qquad (14.118)$$

если обозначить

$$\sum_{x'} \pi(x, y; x', y') = \pi(x, y; y'),$$

и формула (14.114) принимает вид

$$H_{y_k \mid y_1 \ldots y_{k-1}} = -\mathbf{M} \ln \left[\sum_x W_{k-1}(x)\pi(x, y_{k-1}; y_k)\right]$$

Здесь символ **М** указывает усреднение по $y_1, \ldots, y_{k-1}, y_k$. Если же подставить (14.118) в формулу

$$H_{y_k \mid y_1 \ldots y_{k-1}} = -\mathbf{M} \sum_{y_k} P(y_k \mid y_1, \ldots, y_{k-1}) \ln P(y_k \mid y_1, \ldots, y_{k-1}),$$

то будем иметь

$$H_{y_k \mid y_1 \ldots y_{k-1}} = -\mathbf{M} \sum_{y_k} \left[\sum_x W_{k-1}(x)\pi(x, y_{k-1}; y_k)\right] \times \ln\left[\sum_x W_{k-1}(x)\pi(x, y_{k-1}; y_k)\right], \qquad (14.119)$$

где символ **М** обозначает усреднение по $y_1, \ldots, y_{k-1}$.

Удобно рассматривать апостериорные вероятности $W_{k-1}(\bullet)$ как распределение по обеим переменным $(x, y) = \xi$, доопределяя их формулой

$$W_{k-1}(x,\,y)=\begin{cases}W_k(x) & \text{при } y=y_{k-1},\\ 0 & \text{при } y\neq y_{k-1}.\end{cases} \qquad (14.120)$$

Тогда формула (14.118) примет вид

$$P(y_k\,|\,y_1\ldots y_{k-1})=\sum_{\xi}W_{k-1}(\xi)\,\pi(\xi,\,y_k) \qquad (14.120а)$$

и равенство (14.119) можно будет записать

$$H_{y_k\,|\,y_1\ldots y_{k-1}}=$$
$$=-\mathbf{M}\sum_{y_k}\sum_{\xi}W_{k-1}(\xi)\,\pi(\xi,\,y_k)\ln\sum_{\xi'}W_{k-1}(\xi')\,\pi(\xi',\,y_k). \qquad (14.121)$$

Здесь, как и в (14.119), **М** указывает усреднение по $y_k, \ldots, y_{k-1,}$ но подлежащее усреднению выражение, очевидно, зависит от $y_k, \ldots, y_{k-1}$ лишь через посредство апостериорных вероятностей W_{k-1} (•). Поэтому в (14.121) можно считать, что символ **М** соответствует усреднению по случайным величинам $W_{k-1}(A_1), \ldots, W_{k-1}(A_L), (A_1, \ldots, A_L$ — пространство состояний процесса $\xi = (x, y)$, L — число его состояний). Вводя закон распределения $P(dW_{k-1})$ для этих случайных величин, придаем формуле (14.121) вид

$$H_{y_k\,|\,y_1\ldots y_{k-1}}=$$
$$=-\int P(dW_{k-1})\sum_{y'}\sum_{\xi}W_{k-1}(\xi)\,\pi(\xi,\,y')\ln\sum_{\xi'}W_{k-1}(\xi')\,\pi(\xi',\,y'). \qquad (14.122)$$

Это и есть основной результат. Мы видим, что для вычисления условной энтропии части компонент марковского процесса нужно исследовать апостериорные вероятности $\{W_k$ (•)$\}$ как самостоятельный случайный процесс. Этот процесс изучается в теории условных марковских процессов. Известно, что с ростом k указанные вероятности преобразуются по определенным рекуррентным соотношениям. Чтобы вывести последние, запишем равенство, аналогичное (14.117), но с заменой k — 1 на k:

$$W_k(x_k)=\frac{\sum\limits_{x_{k-1}}\Big[\sum\limits_{x_1\ldots x_{k-2}}P(\xi_1)\,\pi(\xi_1,\,\xi_2)\ldots\pi(\xi_{k-2},\,\xi_{k-1})\Big]\pi(\xi_{k-1},\,\xi_k)}{\sum\limits_{x_{k-1}\,x_k}\Big[\sum\limits_{x_1\ldots x_{k-2}}P(\xi_1)\,\pi(\xi_1,\,\xi_2)\ldots\pi(\xi_{k-2},\,\xi_{k-1})\Big]\pi(\xi_{k-1},\,\xi_k)}.$$

Подставляя сюда (14.117), получаем рекуррентные соотношения

$$W_k(x_k)=\frac{\sum\limits_{x_{k-1}}W_{k-1}(x_{k-1})\,\pi(x_{k-1},\,y_{k-1};\,x_k,\,y_k)}{\sum\limits_{x_{k-1},\,x_k}W_{k-1}(x_{k-1})\,\pi(x_{k-1},\,y_{k-1};\,x_k,\,y_k)}. \qquad (14.123)$$

Процесс $\{W_k(\bullet)\}$, рассматриваемый как самостоятельный случайный процесс, носит название вторичного апостериорного *W*-процесса. Как известно из теории условных марковских процессов, он является марковским. Рассмотрим его вероятность перехода.

Преобразование (14.123), которое можно также записать

$$W_k(\xi) = \frac{\sum_{\xi'} W_{k-1}(\xi')\,\pi(\xi',\ \xi)\,\delta_{yy_k}}{\sum_{\xi''} W_{k-1}(\xi'')\,\pi(\xi'',\ y_k)}, \tag{14.124}$$

явно зависит от случайной величины y_k. Ее апостериорные вероятности $P(y_k|y_1,\ ...,\ y_{k-1})$ совпадают с (14.118) и, следовательно, полностью определяются «состоянием» $\{W_{k-1}(\bullet)\}$вторичного апостериорного процесса в момент k — 1. Итак, каждое преобразование (14.124) осуществляется с вероятностью (14.118). Это значит, что «плотность вероятности» перехода вторичного процесса $\{W_k\}$ можно записать

$$\pi^W(W_{k-1},\ W_k) = \sum_{y_k} \delta\left(W_k(\xi) - \frac{\delta_{yy_k}\sum_{\xi'} W_{k-1}(\xi')\,\pi(\xi',\ \xi)}{\sum_{\xi''} W_{k-1}(\xi'')\,\pi(\xi'',\ y_k)}\right) \times$$
$$\times \sum_{\xi'} W_{k-1}(\xi')\,\pi(\xi',\ y_k). \tag{14.125}$$

Здесь $\delta(W)$ — L-мерная дельта-функция, определяемая формулой

$$\int f(W)\,\delta(W) \prod_{\xi} dW(\xi) = f(0).$$

Она соответствует мере, сосредоточенной в начале координат L-мерного пространства.

Итак, мы видим, что *энтропия процесса $\{y_k\}$, немарковского, но являющегося частью марковского процесса, может быть определена методами теории марковских процессов* после перехода к вторичному апостериорному уже марковскому процессу. В стационарном случае обычным способом может быть найдено стационарное распределение $P_{CT}(dW)$, к которому сходится при $k\to\infty$ распределение $P_k(dW)$, входящее в (14.122). Подставляя (14.122) в (14.115) и переходя к пределу $k\to\infty$, получаем удельную энтропию

$$h_y = -\int P_{\text{ст}}(dW) \sum_{y\xi} W(\xi)\,\pi(\xi,\ y) \ln \sum_{\xi'} W(\xi')\,\pi(\xi',\ y). \tag{14.126}$$

Затем уже по формуле (14.107), если нужно, можно найти $h_{x|y}$.

Прежде чем перейти к рассмотрению примера, приведем следующую теорему, имеющую прямое отношение к вышеизложенному.

Теорема 13. *Энтропия* $H_{y_2 \ldots y_l \mid y_1}$ *немарковского у-процесса совпадает с аналогичной энтропией вторичного апостериорного (марковского) процесса:*

$$H_{y_2 \ldots y_l \mid y_1} = H_{W_2 \ldots W_l \mid W_1}. \tag{14.127}$$

Следовательно, совпадают и удельные энтропии: $h_y = h_W$.

Доказательство. Поскольку справедливы равенства

$$H_{y_2 \ldots y_l \mid y_1} = \sum_{k=2}^{l} H_{y_k \mid y_1 \ldots y_{k-1}},$$

$$H_{W_2 \ldots W_l \mid W_1} = \sum_{k=2}^{l} H_{W_k \mid W_1 \ldots W_{k-1}},$$

а также формула (14.115) и аналогичная формула для $\{W_k\}$, то достаточно доказать равенство

$$H_{y_k \mid y_1 \ldots y_{k-1}} = H_{W_k \mid W_1 \ldots W_{k-1}} = H_{W_k \mid W_{k-1}}. \tag{14.128}$$

Здесь

$$H_{W_k \mid W_1 \ldots W_{k-1}} = H_{W_k \mid W_{k-1}}$$

в силу марковского свойства. Пусть S_0 — некоторая исходная точка пространства значений $W(\bullet)$. В соответствии с (14.124) из нее возможны переходы в различные точки $S(y_k)$ в зависимости от значения y_k ($S(y_k)$ — точка с координатами (14.124)). Эти точки не совпадают при несовпадающих y_k. В самом деле, если бы две точки из $S(y_h)$, скажем $S' = S(y')$ и $S'' = S(y'')$, совпадали, т. е. имело бы место равенство

$$W'_k(x, y) = W''_k(x, y),$$

из него можно было бы получить

$$\sum_x W'_k(x, y) = \sum_x W''_k(x, y). \tag{14.129}$$

Но согласно (14.120)

$$\sum_x W'_k(x, y) = \delta_{yy_k} = \delta_{yy'},$$

и равенство (14.120) означало бы совпадение $y' = y''$.

Итак, между точками $S(y_k)$ и значениями y_k можно установить взаимно однозначное соответствие. Поэтому

$$P(S(y_k) \mid S_0) = P(y_k \mid y_1, \ldots, y_{k-1}),$$

$$H_{S(y_k)}(\mid S_0) = H_{y_k}(\mid y_1, \ldots, y_{k-1}).$$

Но при должном выборе точки $S_0 = W_{k-1}$ события, стоящие в условии, совпадают, и после усреднения мы получаем (14.128). Это завершает доказательство.

3. Пример. Пусть немарковский процесс $\{y_k\}$ есть процесс с двумя состояниями, т. е. y_k может принимать одно из двух значений, скажем 1 или 2. Пусть далее его можно превратить в марковский, если добавлением переменной x разбить состояние $y=2$ на два состояния: $\xi= 2$ и состояние $\xi = 3$. Именно, $\xi = 2$ соответствует $x = 1$, $y = 2$, а $\xi = 3$ соответствует $x = 2$, $y = 2$. С состоянием $y = 1$ можно сопоставить, скажем, $\xi = 1$ и $x = 1$.

Совокупный процесс $\{\xi_k\}=\{x_k, y_k)$ является стационарным марковским и описывается матрицей перехода

$$\pi_{\xi\xi'} = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix}.$$

Апостериорные распределения $W_k(\xi) = W_k(x, y)$ согласно (14.120) имеют вид

$$\begin{aligned} &(W_k(\xi=1),\ W_k(\xi=2),\ W_k(\xi=3)) = (1,\ 0,\ 0) \quad \text{при } y_k = 1, \\ &W_k(\xi=1),\ W_k(\xi-2),\ W_k(\xi=3)) = (0,\ W_k(2),\ W_k(3)) \quad \text{при } y_k = 2. \end{aligned} \tag{14.130}$$

Значению $y_k = 2$ согласно (14.124) соответствует преобразование

$$\begin{aligned} &W_k(1) = 0; \\ &W_k(2) = \frac{W_{k-1}(1)\,\pi_{12} + W_{k-1}(2)\,\pi_{22} + W_{k-1}(3)\,\pi_{32}}{W_{k-1}(1)\,(\pi_{12}+\pi_{13}) + W_{k_1}(2)\,(\pi_{22}+\pi_{23}) + W_{k-1}(3)\,(\pi_{32}+\pi_{33})}; \\ &W_k(3) = 1 - W_k(2). \end{aligned} \tag{14.131}$$

Точку (1, 0, 0) трехмерного пространства значений $(W(1), W(2), W(3))$, соответствующую распределению (14.130), обозначим S_0. Исследуем возможные переходы из этой точки. Подставляя значение $W_{k-1} =(0, 0, 0)$ в (14.131), получаем переход в точку

$$\begin{aligned} &W(0) = 0; \\ &W(2) = \frac{\pi_{12}}{\pi_{12}+\pi_{13}}; \\ &W(3) = \frac{\pi_{13}}{\pi_{12}+\pi_{13}}, \end{aligned} \tag{14.132}$$

которую обозначим S_1. Вследствие (14.121) такой переход осуществляется с вероятностью

$$p_1 = \pi_{12} + \pi_{13}. \tag{14.133}$$

С вероятностью $1 — p_1 = \pi_{11}$ сохраняется пребывание в точке S_0. Рассмотрим теперь переходы из точки S_1. Подставляя значения (14.132) в качестве W_{k-1} в формулу (14.131) при $y_k = 2$, получаем координаты

$$\begin{aligned}&W(1)=0;\\&W(2)=\frac{\pi_{12}\pi_{22}+\pi_{13}\pi_{32}}{\pi_{12}(\pi_{22}+\pi_{23})+\pi_{13}(\pi_{32}+\pi_{33})};\\&W(3)=1-W(2)\end{aligned} \quad (14.134)$$

новой точки S_2. Переход из S_2 в S_2 происходит с вероятностью

$$p_2=\frac{\pi_{12}(\pi_{22}+\pi_{23})+\pi_{13}(\pi_{32}+\pi_{33})}{\pi_{12}+\pi_{13}}. \quad (14.135)$$

Это выражение получено подстановкой (14.132) в формулу

$$\begin{aligned}&\mathbf{P}(y_k=2\mid y_1,\ \dots,\ y_{k-1})=P(S_k\mid S_{k-1})=\\&=W_{k-1}(2)(\pi_{22}+\pi_{23})+W_{k-1}(3)(\pi_{32}+\pi_{33}),\end{aligned} \quad (14.136)$$

получаемую из (14.120а). С вероятностью $1—p_2$ происходит возврат в точку S_0. Аналогично, подставляя значения (14.134) в (14.136), находим вероятность

$$p_3=\frac{(\pi_{12}\pi_{22}+\pi_{13}\pi_{32})(\pi_{22}+\pi_{23})+(\pi_{12}\pi_{23}+\pi_{13}\pi_{33})(\pi_{32}+\pi_{33})}{\pi_{12}(\pi_{22}+\pi_{23})+\pi_{13}(\pi_{32}+\pi_{33})} \quad (14.137)$$

перехода из S_2 в следующую точку S_3 и т. д. Описанным способом последовательно вычисляются вероятности перехода p_k в следующие друг за другом точки S_h. Каждый раз с вероятностью $1 — p_k$ происходит возврат в точку S_0. Вероятность того, что к моменту k еще не произошел возврат в точку S_0, очевидно, равна $p_1p_2 \dots p_k$. Если последовательные значения p_k не стремятся к единице, то эта вероятность стремится к нулю при $k\to\infty$. Поэтому обычно с вероятностью, равной единице, происходит в итоге возврат в точку S_0. Если бы в качестве начальной точки была взята какая-либо другая точка S_0, то происходило бы движение по какой-то другой последовательности точек с последующим переходом (сколь угодно близким к достоверности) в точку S_0. После такого перехода будет происходить уже описанное выше движение по точкам S_0, S_1, S_2 , ... с вычисленными вероятностями переходов.

Описанная выше картина переходов вторичного марковского процесса позволяет легко найти стационарное распределение вероятностей. Оно будет сосредоточено в точках S_0, S_1, S_2, ...

Для этих точек матрица вероятностей перехода имеет вид

$$\pi^W=\begin{pmatrix}1-p_1 & p_1 & 0 & 0 & .\\ 1-p_2 & 0 & p_2 & 0 & .\\ 1-p_3 & 0 & 0 & p_3 & .\\ . & . & . & . & .\end{pmatrix}. \quad (14.138)$$

Учитывая равенства $P_{ст}(S_k) = p_k P_{ст}(S_{k-1}),\ k = 1, 2, \ldots,$ вытекающие из (14.87), (14.138), и условие нормировки, находим стационарные вероятности указанных точек

$$P_{ст}(S_0) = \frac{1}{1+p_1+p_1 p_2+\ldots}; \quad P_{ст}(S_k) = \frac{p_1 \ldots p_k}{1+p_1+p_1 p_2+\ldots},$$
$$k = 1, 2, \ldots \tag{14.139}$$

Для вычисления удельной энтропии (14.115) остается лишь подставить выражения (14.139) в формулу

$$P_{ст}(S_0) = \frac{1}{1+p_1+p_1 p_2+\ldots}; \quad P_{ст}(S_k) = \frac{p_1 \ldots p_k}{1+p_1+p_1 p_2+\ldots},$$
$$k = 1, 2, \ldots \tag{14.140}$$

которая вытекает из (14.126) или из теоремы 13 и формулы (14.88), примененной к вероятностям перехода (14.138).

В том частном случае, когда значение x не влияет на переходы от одного значения y к другому, имеем

$$\pi_{22} + \pi_{23} = \pi_{32} + \pi_{33}. \tag{14.141}$$

Здесь выписана вероятность

$$\mathbf{P}(y_k = 2 \mid y_{k-1} = 2,\ x) = 1 - \nu$$

(ν - вероятность перехода $\mathbf{P}(y_k = 1 | y_{k-1} = 2, x)$), которая теперь не зависит от x. В этом случае (14.136) дает

$$p_2 = p_3 = \ldots = [W_{k-1}(2) + W_{k-1}(3)](\pi_{22} + \pi_{23}) =$$
$$= \pi_{22} + \pi_{23} = 1 - \nu \tag{14.142}$$

и из (14.140) получаем

$$h_y = P_{ст}(S_0) h_2(\mu) + [1 - P_{ст}(S_0)] h_2(\nu) \qquad (\mu = p_1 = \pi_{12} + \pi_{13}), \tag{14.143}$$

причем в силу (14.139), (14.142)

$$P_{ст}(S_0) = \left[1 + \frac{p_1}{1-p_2}\right]^{-1} = \frac{\nu}{\nu+\mu}.$$

Формула (14.143) поэтому совпадает с (14.100), и это естественно, поскольку при выполнении условия (14.141) процесс $\{y_k\}$ становится марковским сам по себе.

В рассмотренном примере стационарные вероятности были сосредоточены на счетном множестве точек. Применяя терминологию из п. 14.6, можно сказать, что активное пространство E_a W-процесса состояло из точек $S_0, S_1, S_2, \ldots$ и, следовательно, было счетным, хотя полное пространство $E_a + E_0$ представляло собой континуум. Можно ожидать, что и в других случаях дискретных процессов $\{x_k, y_k\}$ положение такое же, то есть в довольно общем случае конечного или счетного числа состояний совокупного процесса $\{x_k, y_k\}$ стационарное распределение $P_{ст}(dW)$ сосредоточено на счетном множестве E_a точек.

14.4. Энтропия гауссовых случайных величин

1. Рассмотрим l гауссовых случайных величин $\xi_1, ..., \xi_l$, которые описываются вектором средних значений $\mathbf{M}\xi_k = m_k$ и невырожденной корреляционной матрицей

$$R_{ij} = \mathbf{M}(\xi_i - m_i)(\xi_j - m_j). \tag{14.146}$$

Как известно, такие случайные величины имеют плотность распределения вероятностей

$$p(\xi_1, \ldots, \xi_l) = (2\pi)^{-l/2} \det{}^{-1/2} \|R_{ij}\| \times \\ \times \exp\left\{-\frac{1}{2} \sum_{i,j=1}^{l} (\xi_i - m_i) a_{ij} (\xi_j - m_j)\right\}, \tag{14.147}$$

где $\|a_{ij}\| = \|R_{ij}\|^{-1}$ — матрица, обратная корреляционной.

Чтобы определить энтропию указанных величин, введем вспомогательную меру $\nu(d_{\xi_1}, ..., d_{\xi_l})$, удовлетворяющую условию мультипликативности. Для простоты выберем меры $\nu_j(d_{\xi_j})$, которым соответствует одинаковая равномерная плотность распределения

$$\frac{\nu_j(d\xi_j)}{d\xi_j} = \nu_1 = (\nu_0)^{1/l}. \tag{14.148}$$

Как будет видно из дальнейшего, удобно полагать

$$\nu_1 = (2\pi e)^{-1/2}. \tag{14.149}$$

При этом

$$\nu(d\xi_1, \ldots, d\xi_l) = \nu_1^l d\xi_1 \ldots d\xi_l = (2\pi e)^{-l/2} d\xi_1 \ldots d\xi_l.$$

Учитывая (14.147), можно видеть, что при постоянной плотности

$$\nu(d\xi_1, \ldots, d\xi_l)/d\xi_1 \ldots d\xi_l$$

выполнено условие абсолютной непрерывности меры P относительно меры ν. В самом деле, равенство $\nu(A)=0$ для некоторого множества A, состоящего из точек l-мерного действительного пространства R_l означает, что l-мерный объем множества A равен нулю. Но для таких множеств вероятность $P(A)$ также равна нулю. Указанное условие абсолютной непрерывности не было бы выполнено, если бы корреляционная матрица (14.146) была вырожденной. Поэтому условие невырожденности является существенным.

При описанном выборе мер случайная энтропия (14.48) имеет вид

$$H(\xi_1, \ldots, \xi_l) = \frac{l}{2}\ln 2\pi + l\ln \nu_1 + \frac{1}{2}\ln\det\|R_{ij}\| +$$
$$+\frac{1}{2}\sum_{i,j}^{l}(\xi_i - m_i)\,a_{ij}(\xi_j - m_j). \tag{14.150}$$

Чтобы вычислить энтропию (14.47) остается произвести усреднение этого выражения. Учитывая (14.146), получаем

$$H_{\xi_1,\ldots,\xi_l} = \frac{l}{2}\ln 2\pi + l\ln \nu_1 + \frac{1}{2}\ln\det\|R_{ij}\| +$$
$$+\frac{1}{2}\sum_{i,j=1}^{l} a_{ij}R_{ij} = \frac{l}{2}\ln 2\pi + l\ln \nu_1 +$$
$$+\frac{1}{2}\ln\det\|R_{ij}\| + \frac{l}{2}. \tag{14.151}$$

При выборе (14.149) имеет место простой результат

$$H_{\xi_1\ldots\xi_l} = \frac{1}{2}\ln\det\|R_{ij}\|. \tag{14.152}$$

Матрица R $=\|R_{ij}\|$ является симметричной. Поэтому, как известно, существует унитарное преобразование U, приводящее эту матрицу к диагональному виду:

$$\sum_{i,j} U_{ir}^{*} R_{ij} U_{js} = \lambda_r \delta_{rs}.$$

Здесь λ_r — собственные значения корреляционной матрицы, удовлетворяющие уравнению

$$\sum_{k} R_{jk} U_{kr} = \lambda_r U_{jr}. \tag{14.153}$$

При помощи этих собственных значений энтропию $H_{\xi_1,\ldots,\xi_l}$ можно записать в виде

$$H_{\xi_1,\ldots,\xi_l} = \frac{1}{2}\sum_{r=1}^{l}\ln\lambda_r. \tag{14.154}$$

Этой формуле можно придать также вид

$$H_{\xi_1\ldots\xi_l} = \frac{1}{2}\,\mathrm{Sp}\ln R. \tag{14.155}$$

Полученный результат позволяет, в частности, легко найти условную энтропию

$$H_{\xi_l|\xi_1\ldots\xi_{l-1}}.$$

Используя иерархическое свойство энтропии, можно записать

$$H_{\xi_1, \dots, \xi_k} = H_{\xi_1} + H_{\xi_2 \mid \xi_1} + \dots + H_{\xi_k \mid \xi_1, \dots, \xi_{k-1}},$$
$$H_{\xi_1, \dots, \xi_{k-1}} = H_{\xi_1} + H_{\xi_2 \mid \xi_1} + \dots + H_{\xi_{k-1} \mid \xi_1, \dots, \xi_{k-1}}$$

и, вычтя одно выражение из другого, получить

$$H_{\xi_k \mid \xi_1, \dots, \xi_{k-1}} = H_{\xi_1, \dots, \xi_k} - H_{\xi_1, \dots, \xi_{k-1}}.$$

Каждую из энтропии, входящих в правую часть этого равенства, можно вычислить по формуле (14.154). Это приводит к соотношению

$$H_{\xi_k \mid \xi_1, \dots, \xi_{k-1}} = \frac{1}{2} \operatorname{Sp} \ln r^{(k)} - \frac{1}{2} \operatorname{Sp} \ln r^{(k-1)}, \tag{14.156}$$

где

$$r^{(k)} = \begin{pmatrix} R_{11} & \dots & R_{1k} \\ \dots & \dots & \dots \\ R_{k1} & \dots & R_{kk} \end{pmatrix}, \quad r^{(k-1)} = \begin{pmatrix} R_{11} & \dots & R_{1,k-1} \\ \dots & \dots & \dots \\ R_{k-1,1} & \dots & R_{k-1,k-1} \end{pmatrix}.$$

Аналогичным способом можно вычислить энтропию

$$H_{\xi_k \xi_{k-1} \mid \xi_1, \dots, \xi_{k-2}} = H_{\xi_1, \dots, \xi_k} - H_{\xi_1, \dots, \xi_{k-2}} \text{ и т. п.}$$

2. Выберем теперь более сложные меры $\nu_k(d\xi_k)$. Именно, положим, что они имеют плотность распределения $N_k q_k(\xi_k)$, где $q_k(\xi_k)$ — гауссова плотность распределения

$$q_k(\xi_k) = (2\pi\widetilde{\lambda}_k)^{-1/2} \exp\left\{-\frac{1}{2} \frac{(\xi_k - \widetilde{m}_k)^2}{\widetilde{\lambda}_k}\right\}.$$

Условие мультипликативности предполагается выполненным. При этом случайная энтропия (14.48) оказывается равной

$$H(\xi_1, \dots, \xi_l) = \ln N + \frac{1}{2} \ln \det \| R_{ij} \| - \frac{1}{2} \sum_k \ln \widetilde{\lambda}_k - \frac{1}{2} \sum_k \frac{(\xi_k - \widetilde{m}_k)^2}{\widetilde{\lambda}_k} + \frac{1}{2} \sum_{i,j} (\xi_i - m_i) a_{ij} (\xi_j - m_j), \tag{14.157}$$

где $N = \Pi N_h$.

Ее усреднение приводит теперь к соотношению

$$H_{\xi_1,\ldots,\xi_l}=\ln N+\frac{1}{2}\ln\det\|R_{ij}\|-\frac{1}{2}\sum_k\ln\tilde{\lambda}_k-$$
$$-\frac{1}{2}\sum_k\frac{1}{\tilde{\lambda}_k}\left[R_{kk}+(m_k-\tilde{m}_k)^2\right]+\frac{l}{2}. \qquad (14.158)$$

Вводя матрицу $\tilde{R}=\|\tilde{\lambda}_k\delta_{kr}\|$ и используя матричную форму записи, равенству (14.158) можно придать вид

$$H_{\xi_1,\ldots,\xi_l}=\ln N+\frac{1}{2}\operatorname{Sp}\ln(\tilde{R}^{-1}R)-\frac{1}{2}\operatorname{Sp}\tilde{R}^{-1}(R-\tilde{R})-$$
$$-\frac{1}{2}(m-\tilde{m})^T\tilde{R}^{-1}(m-\tilde{m}). \qquad (14.159)$$

Здесь мы учли, что $\det R-\det\tilde{R}=\det\tilde{R}^{-1}R$; $\operatorname{Sp}\tilde{R}^{-1}\tilde{R}=\operatorname{Sp}1=l$; $m-\tilde{m}$ - матрица-столбец; T означает транспонирование. Согласно (14.159) нетрудно видеть, что мы нашли тем самым энтропию $H^{P/Q}_{\xi_1,\ldots,\xi_l}$ распределения P по распределению Q, которая оказалась равной

$$H^{P/Q}_{\xi_1,\ldots,\xi_l}=\operatorname{Sp}G(\tilde{R}^{-1}R)+\frac{1}{2}(m-\tilde{m})^T\tilde{R}^{-1}(m-\tilde{m}), \qquad (14.160)$$

где $G(x)=(x-1-\ln x)/2$.

Последняя формула получена в предположении мультипликативности меры $v(d\xi_1,\ldots,d\xi_l)$, а, следовательно, и $Q(d\xi_1,\ldots,d\xi_l)$. Однако ее нетрудно распространить и на более общий случай. Пусть мера Q вляется гауссовой и определяется вектором $\tilde{m}_k=M_Q\xi_k$ и корреляционной матрицей $\tilde{R}_{kr}$, необязательно диагональной. Энтропия $H^{p/Q}$ является инвариантной относительно ортогональных преобразований (и вообще не вырожденных линейных преобразований) l-мерного действительного пространства R_l. Совершив преобразование поворота, можно добиться того, чтобы матрица $\tilde{R}$ стала диагональной, и после этого применить формулу (14.160). Но формула (14.160) уже является инвариантной относительно линейных невырожденных преобразований и поэтому справедлива не только в случае диагональной, но и недиагональной матрицы $\tilde{R}$. В самом деле, при линейном преобразовании $\xi'_k=\sum_r C_{kr}\xi_r$, т. е. $\xi'=C\xi$, имеют место преобразования

$$m'-\tilde{m}'=C(m-\tilde{m}),\quad R'=CRC^T,\quad \tilde{R}'=C\tilde{R}C^T$$

и, следовательно,

$$(\bar{R}^{-1})' = (C^T)^{-1}\,\bar{R}^{-1}C^{-1},\ (\bar{R}^{-1}R)' = C^{T^{-1}}\,\bar{R}^{-1}\,RC^T.$$

Поэтому комбинации $\mathrm{Sp}\, f\,(\tilde{R}^{-1}R)$ и $(m - \tilde{m})^T\tilde{R}^{-1}(m - \tilde{m})$ остаются инвариантными. Это доказывает, что формула (14.160) сохраняет свое значение не только в мультипликативном случае, но и в более общем случае, когда формулы (14.154), (14.156), (14.157) могут быть несправедливы.

3. В заключение этого пункта вычислим дисперсию энтропии, которую полезно знать при исследовании вопроса об энтропийной устойчивости семейства гауссовых случайных величин.

Начнем со случайной энтропии (14.150). Вычитая (14.151), находим случайное отклонение

$$H(\xi_1,\ldots,\xi_l) - H_{\xi_1,\ldots,\xi_l} = \frac{1}{2}\sum a_{jk}(\eta_j\eta_k - R_{jk}) =$$

$$= \frac{1}{2}\eta^T a\eta - \frac{1}{2}\mathrm{Sp}\,\mathbf{1} = \frac{1}{2}\eta^T a\eta - \frac{l}{2}, \qquad (14.161)$$

средний квадрат которого совпадает с искомой дисперсией. Мы обозначили $\eta_j = \xi_j - m_j$. При усреднении квадрата этого выражения нужно принять во внимание, что

$$\mathbf{M}\eta_j\eta_k\eta_r\eta_s = R_{jk}R_{rs} + R_{jr}R_{ks} + R_{js}R_{kr} \qquad (14.162)$$

согласно известным свойствам гауссовых величин. Поэтому будем иметь

$$\mathbf{D}H(\xi_1,\ldots,\xi_l) = \mathbf{M}\left[H(\xi_1,\ldots,\xi_l) - H_{\xi_1,\ldots,\xi_l}\right]^2 =$$

$$= \frac{1}{4}\sum a_{jk}\,a_{rs}(R_{jr}R_{ks} + R_{js}R_{kr}) = \frac{1}{2}\mathrm{Sp}\,a\,R\,a\,R, \qquad (14.163)$$

т. е.

$$\mathbf{D}H(\xi_1,\ldots,\xi_l) = \frac{1}{2}\mathrm{Sp}\,\mathbf{1} = \frac{l}{2}.$$

Переходя к случайной энтропии (14.157), имеем

$$H(\xi_1,\ldots,\xi_l) - H_{\xi_1,\ldots,\xi_l} = \frac{1}{2}\eta^T(a - \tilde{R}^{-1})\eta -$$

$$- \frac{1}{2}\mathrm{Sp}\,(a - \tilde{R}^{-1})R - \frac{1}{2}(m - \tilde{m})^T\tilde{R}^{-1}\eta.$$

Используя, в свою очередь, (14.162), получаем

$$\mathbf{D}H(\xi_1,\ldots,\ \xi_l)=\frac{1}{2}\operatorname{Sp}\left[(\mathrm{a}-\widetilde{\mathrm{R}}^{-1})\,\mathrm{R}\right]^2+\frac{1}{4}(\mathrm{m}-\widetilde{\mathrm{m}})^T\times$$
$$\times\widetilde{\mathrm{R}}^{-1}\,R\,\widetilde{\mathrm{R}}^{-1}(\mathrm{m}-\widetilde{\mathrm{m}})=\frac{1}{2}\operatorname{Sp}(1-\widetilde{\mathrm{R}}^{-1}\mathrm{R})^2+$$
$$+\frac{1}{4}(\mathrm{m}-\widetilde{\mathrm{m}})^T\widetilde{\mathrm{R}}^{-1}\mathrm{R}\widetilde{\mathrm{R}}^{-1}(\mathrm{m}-\widetilde{\mathrm{m}}). \qquad (14.164)$$

Не представляет особого труда вычислить и другие статистические характеристики случайной энтропии гауссовых переменных, в частности, ее характеристический потенциал

$$\mu_0(s)=\ln \mathbf{M}e^{sH(\xi_1,\ldots,\xi_l)}$$

Так, подставляя сюда (14.150) и принимая во внимание вид (14.147) плотности распределения вероятностей, получаем

$$\mu_0(s)=\ln\Big\{(2\pi)^{-l/2}(\det{}^{-1/2}\mathrm{R})\exp\left[sH_{\xi_1,\ldots,\xi_l}-\frac{sl}{2}\right]\times$$
$$\times\int\exp\Big\{-\frac{1-s}{2}\sum_{i,\ j}\eta_i\,a_{ij}\,\eta_j\Big\}\,d\eta_1,\ldots,\ d\eta_l\Big\}=$$
$$=sH_{\xi_1,\ldots\xi_l}-\frac{sl}{2}+\ln\det{}^{-1/2}\left[(1-s)\,\mathrm{a}\right]+$$
$$+\ln\det{}^{-1/2}R\quad(\eta_i=\xi_i-m_i). \qquad (14.165)$$

Здесь использована формула

$$\int\exp\left[-\frac{1}{2}\sum_{i,\ j}\eta_i\,d_{ij}\,\eta_j\right]d\eta_1,\ldots,\ d\eta_l=\det{}^{-1/2}\left\|\frac{d_{ij}}{2\pi}\right\|, \qquad (14.166)$$

справедливая при любой невырожденной положительно определенной матрице $\|d_{ij}\|$.

Поскольку а=R^{-1}, то в (14.121) члены с $\ln\ \det^{-1/2}\ \mathrm{R}$ сокращаются, и мы получаем

$$\mu_0(s)=-\frac{l}{2}\ln(1-s)-\frac{sl}{2}+sH_{\xi_1,\ldots,\xi_l},$$

что справедливо при $s<1$. Отсюда можно получить, в частности, формулу (14.163).

14.9. Энтропия стационарной последовательности. Гауссова последовательность

1. В п. 14.5 была рассмотрена энтропия отрезка стационарного процесса $\{\xi_k\}$ в дискретном времени, т. е. стационарной последовательности. При этом предполагалось, что каждый элемент последовательности представляет собой дискретную случайную величину. Приведенное обобщение понятия энтропии позволяет рассмотреть энтропию стационарной последовательности, составленную из произвольных, в том числе непрерывных случайных величин, обобщая тем самым результаты п. 14.5.

Если вспомогательная мера v удовлетворяет условию мультипликативности, т.е когда имеется несколько случайных величин ξ_1, ... ξ_n, удобно выбирать меру v, удовлетворяющую условию полной мультипликативности

$$\nu(d\xi_1,\ldots, d\xi_n) = \prod_{k=1}^{n} \nu_k(d\xi_k),$$

которое мы понимаем как условие согласования вспомогательных мер, то условные энтропии в обобщенной версии обладают всеми свойствами, которыми обладали условные энтропии в дискретной версии. Указанные свойства (и по существу только они) были существенно использованы при изложении материала в п. 14.5. Поэтому все сказанное в п. 14.5 может быть отнесено к произвольным случайным величинам, к энтропии в обобщенной версии.

Мера v предполагается мультипликативной

$$\nu(d\xi_{k_1} \ldots d\xi_{k_r}) = \prod_{i=1}^{r} \nu_{k_i}(d\xi_{k_i}), \qquad (14.167)$$

причем из соображений стационарности «элементарные» меры v_k предполагаются одинаковыми. Предполагается также, что выполнено условие абсолютной непрерывности вероятности $P(d\xi_k)$ относительно $v_k(d\xi_k)$. Процесс $\{\xi_k\}$ является стационарным по отношению к распределению P, т. е. выполнено условие типа (14.167) при любых k_1 ..., k_r, a.

Формулой (14.70) вводится удельная энтропия H_1. Теперь под $H_{\xi_k \mid \xi_{k-l}, \ldots, \xi_{k-1}}$, однако, следует понимать энтропию

$$H_{\xi_k \mid \xi_1, \ldots, \xi_{k-1}} = \mathbf{M} H(\xi_k \mid \xi_1, \ldots, \xi_{k-1}) = = -\ln \frac{P(d\xi_k \mid \xi_1, \ldots, \xi_{k-1})}{\nu_k(d\xi_k)}.$$

так что указанное определение соответствует формуле

$$H_1 = -\int \ln \frac{P(d\xi_k \mid \xi_{k-1}, \xi_{k-2}, \ldots)}{\nu_k(d\xi_k)} P(d\xi_k \mid \xi_{k-1}, \xi_{k-2}, \ldots). \qquad (14.168)$$

В обобщенной версии также справедлива теорема 11, которая доказывается точно так же, как и раньше. Теперь она означает равенство

$$H_1 = -\lim_{l\to\infty} \frac{1}{l} \int \dots \int \ln \frac{P(d\xi_1 \dots d\xi_l)}{\nu_1(d\xi_1)\dots\nu_l(d\xi_l)} P(d\xi_1 \dots d\xi_l). \quad (14.169)$$

Далее, по аналогии с (14.78), (14.79) может быть введена величина

$$2\Gamma = \lim_{n\to\infty} \left[H_{\xi_1\dots\xi_n} - nH_1\right] = \sum_{j=0}^{\infty} \left[H_{\xi_{j+1}|\xi_j\dots\xi_1} - H_1\right], \quad (14.170)$$

которая неотрицательна, поскольку

$$H_{\xi_{j+1}|\xi_j\dots\xi_1} \geqslant H_{\xi_{j+1}|\xi_j\xi_{j-1}\dots} = H_1. \quad (14.171)$$

Она может быть интерпретирована как энтропия концов рассматриваемого отрезка последовательности.

Кроме приведенных выше величин и соотношений, основанных на определении энтропии

$$H_{\xi_k|\xi_1,\dots,\xi_{k-1}} = \mathbf{M}H(\xi_k|\xi_1,\dots,\xi_{k-1}) = \\ = -\ln \frac{P(d\xi_k|\xi_1,\dots,\xi_{k-1})}{\nu_k(d\xi_k)}.$$

могут быть рассмотрены аналогичные величины и соотношения. Именно, по аналогии с (14.168), (14.170) можно ввести

$$H_1^{P/Q} = H^{P/Q}_{\xi_k|\xi_{k-1}\xi_{k-2},\dots} = \int \ln \frac{P(d\xi_k|\xi_{k-1},\xi_{k-2},\dots)}{Q(d\xi_k)} \times \\ \times P(d\xi_k|\xi_{k-1},\xi_{k-2},\dots), \quad (14.172)$$

$$2\Gamma^{P/Q} = \sum_{j=0}^{\infty} \left[H^{P/Q}_{\xi_{j+1}|\xi_j\dots\xi_1} - H_1^{P/Q}\right]. \quad (14.173)$$

Для энтропии $H^{P/Q}$, однако, согласно

$$H_\xi^{P/Q} \leqslant H_{\xi|\eta}^{P/Q}.$$

имеет место неравенство

$$H^{P/Q}_{\xi_{j+1}|\xi_j\dots\xi_1} \leqslant H^{P/Q}_{\xi_{j+1}|\xi_j\dots\xi_1,\dots},$$

обратное неравенству (14.171). Поэтому «энтропия конца» $\Gamma^{p/Q}$ обязана быть неположительной.

Подставляя выражения типа

$$H_{\xi_k|\xi_1,\dots,\xi_{k-1}} = \mathbf{M}H(\xi_k|\xi_1,\dots,\xi_{k-1}) = \\ = -\ln \frac{P(d\xi_k|\xi_1,\dots,\xi_{k-1})}{\nu_k(d\xi_k)}.$$

и (14.168) для условных энтропий, нетрудно убедиться, что разность

$$H_{\xi_{j+1}|\xi_1,\dots,\xi_j} - H_1,$$

а следовательно, и граничная энтропия Г оказывается не зависящей от v_k. Аналогично $\Gamma^{p/Q}$ оказывается не зависящей от Q (при выполнении условия мультипликативности), причем справедливо тождество
$\Gamma^{p/Q}$= - Г. Это полезно иметь в виду при записи формулы (14.80) для обеих энтропии, которая принимает вид

$$H_{\xi_1 \ldots \xi_l} = lH_1 + 2\Gamma + o_l(1), \qquad H^{P/Q}_{\xi_1 \ldots \xi_l} = lH^{P/Q}_1 - 2\Gamma + o_l(1). \tag{14.174}$$

Последние соотношения позволяют найти этропию отрезка стационарного процесса точнее, чем простым умножением удельной энтропии H_1 на его длину l.

2. Найдем удельную энтропию H_1 для случая стационарной гауссовой последовательности.

а) Предположим сначала, что задана стационарная последовательность ξ_1, ... ξ_l на круге, распределение которой инвариантно относительно вращений круга. При этом элементы корреляционной матрицы R_{ij} будут зависеть лишь от разности j — k и удовлетворять условию периодичности: $R_{jk} = R_{j-k}$, $R_{j-k+l} = R_{j-k}$. Уравнение (14.153) при этом будет иметь решения:

$$U_{jr} = \frac{1}{\sqrt{l}} e^{2\pi i j r/l}, \tag{14.175}$$

$$\lambda_r = \sum_{s=0}^{l-1} R_s e^{-2\pi i s r/l}, \qquad r = 0, 1, \ldots, l-1. \tag{14.176}$$

В самом деле, подставляя (14.175) в (14.153), получаем равенство

$$\sum_k R_{j-k} e^{2\pi i k r/l} = \lambda_r e^{2\pi i j r/l},$$

которое удовлетворяется в силу (14.176). Итак (14.175) определяет преобразование, которое диагонализует корреляционную матрицу R_{j-k}. Легко проверить его унитарность. В самом деле, эрмитово-сопряженный оператор

$$U^+ = \| U^*_{jr} \| = \left\| \frac{1}{\sqrt{l}} e^{-2\pi i j r/l} \right\|$$

вследствие равенств

$$\sum_r U_{jr} U^*_{kr} = \frac{1}{l} \sum_{r=1}^{l} e^{2\pi i \frac{(j-k)r}{l}} = \frac{\varepsilon}{l} \sum_{r=0}^{l-1} \varepsilon^r = \frac{\varepsilon}{l} \frac{1-\varepsilon^l}{1-\varepsilon} = \delta_{jk}$$

$$(\varepsilon = e^{2\pi i (j-k)/l})$$

совпадает с обратным оператором U^{-1}.

После вычисления собственных значений (14.175) для получения энтропии $H_{\xi_1, \ldots, \xi_l}$ остается воспользоваться формулой (14.154).

В рассматриваемом случае инвариантности относительно поворотов легко вычислить также энтропию (14.160). Предполагается, конечно, что описанной симметрией («круговой стационарностью») обладает не только мера *P*, но и мера *Q*. Корреляционная матрица $\widetilde{R}_{jk}$ последней обладает, следовательно, теми же свойствами, что и R_{jk}. Кроме того, для обеих мер средние значения m_k, $\widetilde{m}_k$ постоянны (соответственно равны m и $\tilde{m}$).

Унитарное преобразование $U = \| U_{jr} \|$ диагонализирует не только матрицу *R*, но и матрицу $\tilde{R}$ (даже, если условие мультипликативности не выполнено), причем по аналогии с (14.176) ее средние значения имеют вид

$$\tilde{\lambda}_r = \sum_{s=0}^{l-1} \widetilde{R}_s e^{-2\pi i s r/l}, \quad r = 1,\dots, l-1. \quad (14.177)$$

Вектор $m - \tilde{m}$ при этом преобразовании переходит в вектор

$$U^+(m - \tilde{m}) = \frac{1}{\sqrt{l}} \begin{pmatrix} \sum_{j=1}^{l} e^{-2\pi i 0 j/l}(m_j - \tilde{m}_j) \\ \cdots\cdots\cdots\cdots \\ \sum_{j=1}^{l} e^{-2\pi i (l-1) j/l}(m_j - \tilde{m}_j) \end{pmatrix} =$$

$$= \frac{m - \tilde{m}}{\sqrt{l}} \begin{pmatrix} \sum_{j=1}^{l} e^{-2\pi i 0 j/l} \\ \vdots \\ \sum_{j=1}^{l} e^{-2\pi i (l-1) j/l} \end{pmatrix} = \begin{pmatrix} \sqrt{l}\,(m - \tilde{m}) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

В результате в соответствии с формулой (14.160) будем иметь

$$H^{P/Q}_{\xi_1 \dots \xi_l} = \sum_{r=0}^{l-1} G\left(\frac{\lambda_r}{\tilde{\lambda}_r}\right) + \frac{l}{2} \frac{(m - \tilde{m})^2}{\tilde{\lambda}_0}, \quad (14.178)$$

где $G(x) = \frac{1}{2}(x - 1 - \ln x)$.

Для энтропии, отнесенной к одному элементу последовательности, согласно (14.154), (14.178) получаем

$$H_1 = \frac{1}{2l} \sum_{r=0}^{l-1} \ln \lambda_r, \quad (14.179)$$

$$H_1^{P/Q} = \frac{1}{l} \sum_{r=0}^{l-1} \ln\left(\frac{\lambda_r}{\tilde{\lambda}_r}\right) + \frac{1}{2} \frac{(m-\tilde{m})^2}{\tilde{\lambda}_0}. \qquad (14.180)$$

б) Пусть теперь $\xi_1, \ldots, \xi_l$ составляют отрезок стационарной последовательности, так что $R_{jk} = R_{j-k}$, $j, k = 1, \ldots, l$, но условие периодичности $R_{j-k+l} = R_{j-k}$ не выполняется. Тогда каждый из концов отрезка играет некоторую роль и дает свой вклад Г в суммарную энтропию (14.174). Если этим вкладом пренебречь, то можно соединить концы отрезка и перейти к разобранному ранее случаю круговой симметрии. При этом нужно образовать новую корреляционную функцию

$$\bar{R}_{j,\,j+s} = \tilde{R}_s = \sum_{n=-\infty}^{\infty} R_{s+nl}, \qquad (14.181)$$

которая уже обладает свойством периодичности. Если R_s заметно отличны от нуля лишь при $\| s \| \ll l$, то добавочные члены в (14.181) заметно повлияют лишь на небольшое число элементов корреляционной матршсы $\tilde{R}_{jk}$, стоящих в углах, где $j \ll l$, $l - k \ll l$ или $l - j \ll l$, $k \ll l$.

После перехода к корреляционной матрице (14.181) (и, если нужно, после аналогичного перехода для второй матрицы $\tilde{R}$) можно воспользоваться полученными ранее формулами (14.179), (14.180), (14.176), (14.177).

Учитывая (14.176), для собственных значений теперь будем иметь

$$\lambda_r = \sum_{n=-\infty}^{\infty} \sum_{s=0}^{l-1} R_{s+nl}\, e^{-2\pi i s r/l} = \sum_{\sigma=-\infty}^{\infty} R_\sigma\, e^{-2\pi i \sigma r/l},$$

или $\lambda_r = \varphi\,(r/l)$, если ввести обозначение

$$\varphi(\mu) = \sum_{\sigma=-\infty}^{\infty} R_\sigma\, e^{-2\pi i \mu\sigma}. \qquad (14.182)$$

Аналогично

$$\tilde{\lambda}_r = \tilde{\varphi}\left(\frac{r}{l}\right) = \sum_{\sigma=-\infty}^{\infty} e^{-2\pi i \sigma r/l}\, \tilde{R}_\sigma.$$

Формулы (14.179), (14.180), очевидно, примут вид

$$\bar{H}_1 = \frac{1}{2l} \sum_{r=0}^{l-1} \ln \varphi\left(\frac{r}{l}\right), \qquad (14.183)$$

$$\bar{H}_1^{P/Q} = \frac{1}{l} \sum_{r=0}^{l-1} G\left(\frac{\varphi(r/l)}{\tilde{\varphi}(r/l)}\right) + \frac{1}{2} \frac{(m-\tilde{m})^2}{\tilde{\varphi}(0)}.$$

Будем теперь увеличивать длину l выбранного отрезка последовательности. Совершая в последних формулах предельный переход $l \to \infty$, получаем вместо сумм интегралы

$$H_1 = \frac{1}{2}\int_0^1 \ln \varphi(\mu)\, d\mu = \frac{1}{2}\int_{-1/2}^{1/2} \ln \varphi(\mu)\, d\mu. \qquad (14.184)$$

$$H_1^{P/Q} = \int_0^1 G\left(\frac{\varphi(\mu)}{\widetilde{\varphi}(\mu)}\right) d\mu + \frac{1}{2}\frac{(m-\widetilde{m})^2}{\widetilde{\varphi}(0)} =$$

$$= \int_{-1/2}^{1/2} G\left(\frac{\varphi(\mu)}{\widetilde{\varphi}(\mu)}\right) d\mu + \frac{(m-\widetilde{m})^2}{2\widetilde{\varphi}(0)}. \qquad (14.185)$$

Здесь при изменении пределов интегрирования учтено свойство $\varphi(\mu + 1) = \varphi(\mu)$, вытекающее из (14.182).

При больших l концы отрезка дают относительно малый вклад по сравнению с большой полной энтропией, имеющей порядок lH_1. Переход (14.181) к корреляционной функции $\overline{R}_s$ изменяет энтропию $H_{\xi_1,\ldots,\xi_l}$ на некоторое число, не возрастающее с ростом l. Поэтому имеет место совпадение пределов

$$\lim_{l\to\infty} \frac{1}{l}\overline{H}_{\xi_1,\ldots,\xi_l} = \lim_{l\to\infty}\frac{1}{l} H_{\xi_1,\ldots,\xi_l}.$$

Здесь $\overline{H}$ соответствует корреляционной функции $\overline{R}_s$, а H —функции R_s. Следовательно, выражения (14.184), (14.185) совпадают с определенными ранее удельными энтропиями (14.169), (14.172). Тем самым мы вычислили удельные энтропии для случая стационарных гауссовых последовательностей. Они оказались выраженными через спектральные плотности $\varphi(\mu), \widetilde{\varphi}(\mu)$.

Условие абсолютной непрерывности меры P относительно меры Q принимает вид условия интегрируемости функции $G(\varphi(\mu)/\widetilde{\varphi}(\mu))$, входящей в (14.185).

Формула (14.185) справедлива не только при выполнении условия мультипликативности (14.167). Для стационарного гауссового случая это условие означает, что матрица $\widetilde{R}_{jk}$ кратна единичной:

$$\widetilde{R}_{jk} = \widetilde{\varphi}\delta_{jk}, \quad \widetilde{\varphi}(\mu) = \widetilde{\varphi} = \mathrm{const}.$$

Тогда формула (14.185) дает

$$H_1^{P/Q} = \int_0^1 G\left(\frac{\varphi(\mu)}{\widetilde{\varphi}}\right) d\mu + \frac{1}{2}\frac{(m-\widetilde{m})^2}{\widetilde{\varphi}}.$$

Приведенные результаты обобщаются и на тот случай, когда имеется не одна случайная последовательность $\{..., \xi_1, \xi_2...\}$, а несколько *(r)* стационарных и стационарно связанных последовательностей $\{\ldots, \xi_1^\alpha, \xi_2^\alpha, \ldots\}$, $\alpha = 1, \ldots, r$, описываемых корреляционной матрицей

$$\| R_{j-k}^{\alpha, \beta} \|, \qquad \alpha, \beta = 1, \ldots, r$$

или матрицей спектральных функций

$$\varphi(\mu) = \| \varphi^{\alpha\beta}(\mu) \|, \; \varphi^{\alpha\beta}(\mu) = \sum_{\sigma=-\infty}^{\infty} R_\sigma^{\alpha\beta} e^{-2\pi i \mu\sigma}$$

и вектором-столбцом (по индексу α) средних значений $m = \| m^n \|$.

При этом формула (14.184) заменяется на матричное обобщение

$$H_1 = \frac{1}{2} \int_{-1/2}^{1/2} \operatorname{Sp} \left[\ln \varphi(\mu)\right] d\mu = \frac{1}{2} \int_{-1/2}^{1/2} \ln \left[\det \varphi(\mu)\right] d\mu. \quad (14.186)$$

Мера *Q* пусть описывается матрицей спектральных функций $\tilde{\varphi}(\mu) = \| \varphi^{\alpha\beta}(\mu) \|$ и средних значений $\tilde{m} = \| \tilde{m}^\alpha \|$. Тогда, вместо (14.185), будем иметь аналогичную матричную формулу

$$H_1^{P/Q} = \int_{-1/2}^{1/2} \operatorname{Sp} G\left(\tilde{\varphi}^{-1}(\mu)\varphi(\mu)\right) d\mu + \frac{1}{2}(m - \tilde{m})^T \tilde{\varphi}^{-1}(0)(m - \tilde{m}).$$

(14.187)

Приведенные результаты вытекают из формул (14.152), (14.160). По виду они представляют собой синтез формул (14.152), (14.160) и (14.184), (14.185).

3. Полученные результаты позволяют сделать заключение об энтропийной устойчивости семейства случайных величин $\{\xi^l\}$, где $\xi^l = \{\xi_1, \ldots, \xi_l\}$ — отрезок стационарной гауссовой последовательности. Энтропия $H_{\xi_1, \ldots, \xi_l}$ растет с ростом *l* приблизительно линейно. Дисперсия энтропии согласно (14.163) также растет линейно. Поэтому отношение $\mathbf{D}\, H(\xi_1, \ldots, \xi_l)/H^2_{\xi_1, \ldots, \xi_l}$ стремится к нулю, так что условие энтропийной устойчивости для энтропии (14.150) оказывается выполненным.

Перейдем к энтропии (14.157). Условия

$$\frac{\mathbf{D} H(\xi_1, \ldots, \xi_l)}{H^2_{\xi_1, \ldots, \xi_l}} \to 0, \; \frac{\mathbf{D} H^{P/Q}(\xi_1, \ldots, \xi_l)}{\left(H^{P/Q}_{\xi_1, \ldots, \xi_l}\right)^2} \to 0$$

для нее будут выполнены, если дисперсия

$$\mathbf{D}H\ (\xi_1,\ \ldots,\ \xi_l) = \mathbf{D}H^{P/Q}\ (\xi_1,\ \ldots,\ \xi_l),$$

определяемая формулой (14.164), возрастает с ростом l приблизительно линейно, т. е. если существует конечный предел

$$D_1 = \lim_{l\to\infty} \frac{1}{l} \mathbf{D}H^{P/Q}(\xi_1,\ldots,\ \xi_l). \tag{14.188}$$

Для вычисления предела (14.188) выражения (14.164) можно применить те же методы, какие были использованы для вычисления предела

$$\lim_{l\to\infty} \frac{1}{l}\ H^{P/Q}_{\xi_1,\ \ldots,\ \xi_l},$$

соответствующего выражению (14.160).

Подобно тому, как из (14.160) мы получили (14.185), из формулы (14.164) находим предел (14.188)

$$D_1 = \frac{1}{2} \int_{-1/2}^{1/2} \left[1 - \frac{\varphi(\mu)}{\tilde{\varphi}(\mu)} \right]^2 d\mu + \frac{(m-\tilde{m})^2\, \varphi(0)}{4\tilde{\varphi}^2(0)}\ . \tag{14.189}$$

14.10. Энтропия случайных процессов в непрерывном времени.

1. Обобщенное определение энтропии, позволяет вычислять энтропию случайных процессов $\{\xi_t\}$, зависящих от непрерывного параметра t (времени). Будем предполагать, что процесс ξ_t задан на некотором интервале $a \le t \le b$. Для произвольного подынтервала $\alpha \le t \le \beta$, принадлежащего интервалу определения процесса, будем пользоваться обозначением $\xi_\alpha^\beta = \{\xi_t,\ \alpha \leqslant t \leqslant \beta\}$. Следовательно, ξ_α^β обозначает совокупность значений процесса $\{\xi_t\}$ на подынтервале $[\alpha,\beta]$.

Исходный процесс $\{\xi_t\}$ описывается вероятностной мерой P. В соответствии с определением энтропии, чтобы определить энтропию $H_{\xi_\alpha^\beta}$ для любых различных интервалов $[\alpha,\ \beta]$, нужно ввести вспомогательную ненормированную меру ν или соответствующую ей вероятностную меру Q. Мера ν (или Q) должна быть определена на том же измеримом пространстве, т. е. на том же поле событий, касающихся поведения процесса $\xi(f)$ на всем интервале $[a,\ b]$. При этом процесс $\{\xi(t)\}$, имеющий вероятности Q, можно интерпретировать как новый вспомогательный случайный процесс $\{\eta(t)\}$, отличающийся от исходного процесса $\{\xi\ (t)\}$.

Мера *P* должна быть абсолютно непрерывна относительно меры *Q* (или *v*) для всего поля событий, относящихся к поведению процесса $\xi(t)$ на всем интервале определения [*a, b*]. Следовательно, условие абсолютной непрерывности будет выполнено и для любого его подынтервала [α, β].

Применяя формулу (14.51) для значений случайного процесса на некотором выбранном подынтервале [α, β], получаем следующее определение энтропии для этого подынтервала:

$$H^{P/Q}_{\xi_\alpha^\beta} = \int \ln \frac{P\left(d\xi_\alpha^\beta\right)}{Q\left(d\xi_\alpha^\beta\right)} P\left(d\xi_\alpha^\beta\right). \qquad (14.190)$$

Далее можно ввести условную энтропию

$$H^{P/Q}_{\xi_\alpha^\beta \mid \xi_\gamma^\delta} = \int \ln \frac{P\left(d\xi_\alpha^\beta \mid \xi_\gamma^\delta\right)}{Q\left(d\xi_\alpha^\beta\right)} P\left(d\xi_\alpha^\beta\, d\xi_\gamma^\delta\right), \qquad (14.191)$$

где [γ, δ] — другой подынтервал, не перекрывающийся с [α, β].

Введенные энтропии подчиняются обычным соотношениям, встречающимся в дискретной версии, например, соотношению аддитивности

$$H^{P/Q}_{\xi_\alpha^\delta} = H^{P/Q}_{\xi_\alpha^\beta} + H^{P/Q}_{\xi_\beta^\delta \mid \xi_\alpha^\beta}, \quad \alpha < \beta < \delta. \qquad (14.192)$$

При записи формул (14.191), (14.192) предполагается, что меры *Q, v* удовлетворяют условию мультипликативности

$$Q\left(d\xi_\alpha^\beta\, d\xi_\gamma^\delta\right) = Q\left(d\xi_\alpha^\beta\right) Q\left(d\xi_\gamma^\delta\right),\ \nu\left(d\xi_\alpha^\beta\, d\xi_\gamma^\delta\right) = \nu_1\left(d\xi_\alpha^\beta\right) \nu_2\left(d\xi_\gamma^\delta\right) \qquad (14.193)$$

([α, β] не перекрывается с [γ, δ]).

Приведендое условие мультипликативности для меры *Q* означает, что вспомогательный процесс $\{\eta_t\}$ таков, что его значения ξ_α^β и ξ_γ^δ для неперекрывающихся интервалов [α, β], [γ, δ] должны быть независимыми. Условие мультипликативности для меры *v* означает кроме того, что постоянные

$$N_\alpha^\beta = \int \nu\left(d\xi_\alpha^\beta\right)$$

определяются некоторой возрастающей функцией *F (t)* по формуле

$$N_\alpha^\beta = e^{F(\beta) - F(\alpha)}. \qquad (14.194)$$

В случае стационарного процесса функция *F (t)* линейна, так что

$$F(\beta) - F(\alpha) = (\beta - \alpha) h_\nu, \quad N_\alpha^\beta = e^{(\beta-\alpha) h_\nu}.$$

Учитывая (14.194), безусловную энтропию типа (14.50) и условную энтропию типа (14.66) можно определить по формулам

$$H_{\xi_\alpha^\beta} = F(\beta) - F(\alpha) - H^{P/Q}_{\xi_\alpha^\beta}, \tag{14.195}$$

$$H_{\xi_\alpha^\beta \mid \xi_\gamma^\delta} = F(\beta) - F(\alpha) - H^{P/Q}_{\xi_\alpha^\beta \mid \xi_\gamma^\delta}, \tag{14.196}$$

где стоящие справа величины определены соотношениями (14.190), (14.191).

2. В дальнейшем будем рассматривать стационарный случайный процесс $\{\xi_t\}$, определенный при всех t. В этом случае вспомогательный процесс $\{\eta_t\}$ естественно выбирать также стационарным.

Ввиду того, что энтропия в обобщенной версии при выполнении условия мультипликативности обладает теми же свойствами, что и энтропия в дискретной версии, на рассматриваемый случай непрерывного времени могут быть перенесены рассуждения и результаты, изложенные ранее применительно к стационарному процессу в дискретном времени.

Вследствие обычных общих свойств энтропии условная энтропия

$$H_{\xi_0^\tau \mid \xi_{-\sigma}^0} \quad (\sigma > 0)$$

монотонно не возрастает с ростом σ.

Отсюдавытекает существование предела

$$\lim_{\sigma\to\infty} H_{\xi_0^\tau \mid \xi_{-\sigma}^0} = H_{\xi_0^\tau \mid \xi_{-\infty}^0}, \tag{14.197}$$

который мы определяем как $H_{\xi_0^\tau \mid \xi_{-\infty}^0}$.

Поскольку в силу общего свойства (14.65)

$$H_{\xi_0^{\tau_1+\tau_2} \mid \xi_{-\sigma}^0} = H_{\xi_0^{\tau_1} \mid \xi_{-\sigma}^0} + H_{\xi_{\tau_1}^{\tau_1+\tau_2} \mid \xi_{-\sigma}^{\tau_1}}, \tag{14.198}$$

а в силу условия стационарности

$$H_{\xi_{\tau_1}^{\tau_1+\tau_2} \mid \xi_{-\sigma}^{\tau_1}} = H_{\xi_0^{\tau_2} \mid \xi_{-\tau_1-\sigma}^0},$$

то, переходя в (14.198) к пределу σ→∞, получаем

$$H_{\xi_0^{\tau_1+\tau_2} \mid \xi_{-\infty}^0} = H_{\xi_0^{\tau_1} \mid \xi_{-\infty}^0} + H_{\xi_0^{\tau_2} \mid \xi_{-\infty}^0}.$$

Следовательно, условная энтропия $H_{\xi_0^\tau \mid \xi_{-\infty}^0}$ линейно зависит от τ.

Соответствующий коэффициент пропорциональности определяем как удельную энтропию:

$$h = \frac{1}{\tau} H_{\xi_0^\tau \mid \xi_{-\infty}^0}. \tag{14.199}$$

Аналогом теоремы 11 в непрерывной версии будет следующая теорема.

Теорема 14. *Если энтропия* $H_{\xi_0^t}$ *конечная, то удельную энтропию* (14.199) *можно вычислять предельным переходом*

$$h = \lim_{t\to\infty} \frac{1}{t} H_{\xi_0^t} . \tag{14.200}$$

Доказательство проводится тем же методом, что и доказательство теоремы 11. Пользуясь свойством аддитивности (14.67), представим энтропию $H_{\xi_0^t}$ в виде

$$H_{\xi_0^t} = H_{\xi_0^\sigma} + H_{\xi_\sigma^{\sigma+1} | \xi_0^\sigma} + H_{\xi_{\sigma+1}^{\sigma+2} | \xi_0^{\sigma+1}} + \ldots + H_{\xi_{t-1}^t | \xi_0^{t-1}} \qquad (\sigma = t - n). \tag{14.291}$$

Вследствие (14.197), (14.199) имеем

$$H_{\xi_{\sigma+k}^{\sigma+k+1} | \xi_0^{\sigma+k}} = h + o_{\sigma+k}(1) \qquad (0 \leqslant k \leqslant t - \sigma - 1), \tag{14.202}$$

где $o_{\sigma+k}(1)$ стремится к 0 при $\sigma\to\infty$. Подставляя (14.202) в (14.201), получаем

$$\frac{1}{t} H_{\xi_0^t} = \frac{1}{\sigma+n} H_{\xi_0^\sigma} + \frac{n}{\sigma+n} h + \frac{n}{\sigma+n} o_\sigma(1). \tag{14.203}$$

Устремим здесь σ и *n* к бесконечности, но так, чтобы $n/\sigma \to \infty$. Поскольку

$$H_{\xi_0^\sigma} = H_{\xi_0^1} + H_{\xi_1^2 | \xi_0^1} + \ldots + H_{\xi_{m-1}^\sigma | \xi_0^{m-1}} \leqslant m H_{\xi_0^1} \; (m - [\sigma] + 1),$$

то при этом член $\frac{1}{\sigma+n} H_{\xi_0^\sigma} \leqslant \frac{m}{\sigma+n} H_{\xi_0^1}$ будет стремиться к 0. Член $\frac{n}{\sigma+n} o_\sigma(1)$ также будет стремиться к 0, а $\frac{n}{\sigma+n} h$— к h, поскольку $n/(\sigma + n) \to 1$. Следовательно, из (14.203) получим требуемое соотношение (14.200). Доказательство закончено.

На случай непрерывного времени обобщаются и те высказывания, которые касаются граничной энтропии Г. По аналогии с (14.77) ее можно определить по формуле

$$2\Gamma = \lim_{\sigma\to\infty,\, \tau\to\infty} \left(H_{\xi_0^\sigma} + H_{\xi_0^\tau} - H_{\xi_0^{\sigma+\tau}} \right) \tag{14.204}$$

и представить в виде

$$\Gamma = \frac{1}{2}\int_0^\infty \left(H_{\xi_0^{d\tau} | \xi_{-\tau}^0} - h d\tau\right), \qquad (14.205)$$

аналогичном (14.79).

При помощи величин h, Г энтропия $H_{\xi_0^t}$ конечного отрезка стационарного процесса выражается по формуле

$$H_{\xi_0^t} = th + 2\Gamma + o_t(1). \qquad (14.206)$$

Из определения (14.204) граничной энтропии Г, если учесть (14.190), нетрудно усмотреть, что она не зависит от выбора меры Q или v.

Для энтропии $H^{p/Q}$ при выполнении условия мультипликативности по аналогии с (14.206) будет справедлива формула

$$H_{\xi_0^t}^{p/Q} = th^{p/Q} - 2\Gamma + o_t(1), \qquad (14.207)$$

где Г — та же величина, что и в (14.206).

14.11. Энтропия гауссового процесса в непрерывном времени

1. Гауссов случайный процесс $\xi(t)$ при непрерывном времени t $(a \le t \le b)$, подобно случайным величинам, рассмотренным ранее, характеризуется вектором средних значений $\mathbf{M}\xi(t)=m(t)$ и корреляционной матрицей

$$R(t, t') = \mathbf{M}\,[\xi(t) - m(t)]\,[\xi(t') - m(t')].$$

Разница лишь в том, что теперь вектор представляет собой заданную на интервале $[a, b]$ функцию от t. Матрица (скажем, R) теперь является функцией двух переменных t, t', определяющей линейное преобразование

$$y(t) = \int_a^b R(t, t')\,x(t')\,dt' \qquad (t \in [a, b])$$

вектора $x(t)$. На этот случай, как известно, могут быть перенесены все основные результаты теории конечномерных векторов. При этом требуется произвести такое тривиальное видоизменение формул, как замена суммы на интеграл и т. п. С описанными видоизменениями на случай непрерывного времени могут быть распространены изложенные ранее методы вычисления энтропии. Результирующие матричные формулы (14.155), (14.160) сохраняют свое значение при новом понимании матриц и векторов.

Указанные выражения теперь не обязаны быть конечными. Условие их конечности связано с условием абсолютной непрерывности меры *P* относительно меры ν или *Q*.

При обобщенном понимании векторов и матриц формулы (14.155), (14.160) справедливы как для конечных, так и для бесконечных интервалов [*a, b*] определения процесса для стационарного и для нестационарного случаев. В дальнейшем мы ограничимся рассмотрением стационарных процессов и вычислим для них удельные энтропии h, $h^{p/Q}$.

Для этой цели может быть применен прием, заключающийся в переходе к периодическому стационарному процессу. Рассматривая на интервале [0, *T*] процесс, имеющий корреляционную функцию $R(t, t') = R(t - t')$, можно сконструировать новую корреляционную функцию

$$\bar{R}(\tau) = \sum_{n=-\infty}^{\infty} R(\tau + nT), \tag{14.208}$$

которая помимо стационарности будет обладать еще свойством периодичности. Формула (14.208) аналогична формуле (14.181). Такую корреляционную функцию имеет процесс

$$\bar{\xi}(t) - \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \xi(t + jT)$$

в пределе при $N \to \infty$.

Стационарная периодическая матрица $\bar{R}(t - t')$ типа (14.208) диагонализуется унитарным преобразованием U, имеющим матрицу

$$U_{tr} = \frac{1}{\sqrt{T}} e^{2\pi i t r/T} \quad t \in [0, T], \quad r = \ldots, -1, 0, 1, 2, \ldots$$

При этом собственные значения λ матрицы $\bar{R}(t - t')$ оказываются такими:

$$\lambda_r = \int_0^T \bar{R}_\tau e^{-2\pi i \tau r/T} d\tau.$$

Если учесть (14.208), то отсюда будем иметь

$$\lambda_r = \int_{-\infty}^{\infty} e^{-2\pi i \tau r/T} R(\tau) d\tau = S\left(2\pi \frac{r}{T}\right), \tag{14.209}$$

где через $S(\omega)$ обозначена спектральная плотность

$$S(\omega)=\int\limits_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau)\,d\tau=S(-\omega) \qquad (14.210)$$

процесса $\xi\,(t)$.

Подставляя (14.209) в (14.155), получаем

$$\overline{H}_{\xi_0^T}=\frac{1}{2}\operatorname{Sp}\ln \mathrm{R}=\frac{1}{2}\sum_r \ln\lambda_r=\frac{1}{2}\sum_r \ln S\left(2\pi\frac{r}{T}\right). \qquad (14.211)$$

Перейдем к энтропии

$$H^{P/Q}_{\xi_0^T},$$

которая находится по формуле (14.160). Эту формулу удобно применить после диагонализации матриц

$$R\ (t-t'),\ \widetilde{R}\ (t-t').$$

Тогда

$$\operatorname{Sp} G\left(\widetilde{\mathrm{R}}^{-1}\mathrm{R}\right)=\sum_r G\left(\frac{\lambda_r}{\widetilde{\lambda}_r}\right), \qquad (14.212)$$

где

$$\widetilde{\lambda}_r=\widetilde{S}\left(2\pi\frac{r}{T}\right),\qquad \widetilde{S}=\int\limits_{-\infty}^{\infty} e^{-i\omega\tau}\widetilde{R}(\tau)\,d\tau \qquad (14.213)$$

по аналогии с (14.209), (14.210). Второй член в правой части (14.160) после диагонализации примет вид

$$\frac{1}{2}\left(\mathrm{m}^T-\widetilde{\mathrm{m}}^T\right)\widetilde{\mathrm{R}}^{-1}\left(\mathrm{m}-\widetilde{\mathrm{m}}\right)=\mathrm{c}^+\,\mathrm{U}^{-1}\,\widetilde{\mathrm{R}}^{-1}\,\mathrm{U}\mathrm{c}=$$
$$=\sum\widetilde{\lambda}_r^{-1}\,c_r\quad\left(\mathrm{c}=\mathrm{U}^+(\mathrm{m}-\widetilde{\mathrm{m}})\right). \qquad (14.214)$$

Поскольку

$$c_r=\int U^+_{tr}\left[m(t)-\widetilde{m}(t)\right]dt=\begin{cases}\sqrt{T}\ (m-\widetilde{m}) & \text{при } r=0,\\ 0 & \text{при } r\neq 0\end{cases}$$

то из (14.160) вследствие (14.212), (14.214) будем иметь

$$\overline{H}^{P/Q}_{\xi_0^T}=\sum_{r=-\infty}^{\infty} G\left(\frac{S(2\pi r/T)}{\widetilde{S}(2\pi r/T)}\right)+T\left(\frac{(m-\widetilde{m})^2}{\widetilde{S}(0)}\right). \qquad (14.215)$$

Сумма по r в правой части (14.215) содержит бесконечное число членов. Чтобы ряд сходился к конечному пределу, необходимо чтобы отношение $S\ (2\pi r/T)/\widetilde{S}\ (2\pi r/T)$ стремилось к 1 при $|r|\to\infty$, так как функция $G\ (x)\ =\ (x\ -\ 1\ -\ \ln\ x)/2$ обращается в нуль (в области положительных значений x) лишь при x = 1. Вблизи точки x = 1 функция ведет себя так:

$$G(x)=\frac{1}{4}(x-1)^2+O\left((x-1)^3\right).$$

Поэтому при выполнении условий

$$\sum_{r=0}^{\infty}\left[\frac{S(2\pi r/T)}{\widetilde{S}(2\pi r/T)}-1\right]^2<\infty, \qquad 0<\frac{S(2\pi r/T)}{\widetilde{S}(2\pi r/T)}<\infty \quad (14.216)$$

энтропия (14.215) оказывается конечной, что свидетельствует об абсолютной непрерывности меры P относительно Q.

2. Перейдем к вычислению удельной энтропии для стационарного процесса, заданного на бесконечной временной оси. Поскольку замена (14.208) корреляционной функции оказывает существенное влияние лишь на граничные эффекты, то энтропии (14.211), (14.215) отличаются от энтропии, соответствующих корреляционной функции $R(\tau)$ величиной порядка 1, а не порядка $T \gg 1$. Поэтому найденные выражения (14.211), (14.215) можно использовать для вычисления удельной энтропии

$$h=\lim_{T\to\infty}\frac{1}{T}H_{\xi_0^T}=\lim_{T\to\infty}\frac{1}{T}\overline{H}_{\xi_0^T},$$

$$h^{P/Q}=\lim_{T\to\infty}\frac{1}{T}H^{P/Q}_{\xi_0^T}=\lim_{T\to\infty}\frac{1}{T}\overline{H}^{P/Q}_{\xi_0^T}.$$

Итак, в полученных формулах следует совершить предельный переход $T\to\infty$. При этом суммы по r обратятся в интеграл. Из (14.211) будем иметь

$$h=\frac{1}{4\pi}\lim_{T\to\infty}\sum_{r=-\infty}^{\infty}\ln S\left(\frac{2\pi r}{T}\right)\frac{2\pi}{T}=\frac{1}{4\pi}\int_{-\infty}^{\infty}\ln S(\omega)\,d\omega, \quad (14.217)$$

а (14.215) даст

$$h^{P/Q}=\frac{1}{2\pi}\int_{-\infty}^{\infty}G\left(\frac{S(\omega)}{\widetilde{S}(\omega)}\right)d\omega+\frac{(m-\widetilde{m})^2}{\widetilde{S}(0)}=$$

$$=\frac{1}{\pi}\int_{0}^{\infty}G\left(\frac{S(\omega)}{\widetilde{S}(\omega)}\right)d\omega+\frac{(m-\widetilde{m})^2}{\widetilde{S}(0)}. \quad (14.218)$$

Эти результаты можно получить также из формул (14.184), (14.185), определяющих удельную энтропию стационарных последовательностей как предельный результат при неограниченном уплотнении точек на временной оси. Выбрав точки $t_k=k\Delta$ из стационарного гауссова процесса в непрерывном времени t, можно получить стационарную гауссову последовательность $\{\xi(t_k)\}$, которая будет иметь корреляционную матрицу

$$R_{jk} = R((j-k)\Delta)$$

и те же средние значения *m.* Сравнивая формулу (14.182) с (14.210), нетрудно видеть, что

$$\varphi\left(\frac{\omega\Delta}{2\pi}\right)\Delta = \sum_{\sigma=-\infty}^{\infty} R(\sigma\Delta)\,e^{-i\omega\sigma\Delta}\,\Delta = \int_{-\infty}^{\infty} R(\tau)\,e^{-i\omega\tau}d\tau + o_\Delta(1),$$

т. е. функции φ(μ), *S*(ω) связаны между собой соотношением

$$\varphi(\mu)\,\Delta = S(\omega) + o_\Delta(1) \quad \text{при} \quad \mu = \omega\Delta/2\pi,\ |\omega| \sim 1. \qquad (14.219)$$

Относя энтропию не к одному элементу последовательности, а к единице времени, имеем

$$h = \lim_{\Delta\to\infty} H_1/\Delta, \qquad h^{P/Q} = \lim_{\Delta\to 0} H_1^{P/Q}/\Delta.$$

Подставляя сюда (14.184), (14.185) при учете равенства (14.219) и аналогичного равенства для $\tilde{\varphi}$, $\tilde{S}$, получаем

$$h = \frac{1}{4\pi}\lim_{\Delta\to 0}\int_{-\pi/\Delta}^{\pi/\Delta}[-\ln\Delta + \ln S(\omega)]\,d\omega, \qquad (14.220)$$

$$h^{P/Q} = \frac{1}{2\pi}\int_{-\infty}^{\infty} G\left(\frac{S(\omega)}{\tilde{S}(\omega)}\right)d\omega + \frac{(m-\tilde{m})^2}{2\tilde{S}(0)}. \qquad (14.221)$$

Формула (14.221) совпадает с (14.218), а (14.220) отличается от (14.217) членом —ln Δ, который может быть отнесен к соответствующим образом подобранной мере *v*. Пользуясь свободой выбора меры *v*, можно представить формулы (14.217), (14. 220) в другой, более удобной форме. Предположим, что спектральная плотность *S*(ω) стремится при ω→∞ к конечному отличному от нуля пределу

$$S(\infty) = \lim_{|\omega|\to\infty} S(\omega).$$

Учитывая, что согласно (14.219) $\varphi(\mu)\,\Delta = S(2\pi\mu/\Delta) + o_\Delta(1)$, будем иметь, следовательно, что при фиксированном μ в процессе предельного перехода Δ→0 значение φ(μ)Δ будет стремиться к *S*(∞). Это предельное значение целесообразно выделить, записав формулу (14.183) в виде

$$\overline{H}_1 = \frac{1}{2}\ln\frac{S(\infty)}{\Delta} + \frac{1}{2l}\sum_{r=0}^{l-1}\ln\frac{\varphi(r/l)\,\Delta}{S(\infty)}. \qquad (14.222)$$

Определим теперь меру *v* не формулой (14.149), а формулой

$$\frac{\nu(d\xi(t_j))}{d\xi(t_j)} = (2\pi e)^{-1/2}(S(\infty)/\Delta)^{-1/2}. \qquad (14.223)$$

Тогда первый член в (14.222) будет отнесен к мере v и вместо (14.222) будем иметь

$$\overline{H}_1 = \frac{1}{2l} \sum_{r=0}^{l-1} \ln \frac{\varphi(r/l)\,\Delta}{S(\infty)}. \qquad (14.224)$$

Соответственно этому формула (14.224) заменится формулой

$$H_1 = \frac{1}{2} \int_{-1/2}^{1/2} \ln \frac{\varphi(\mu)\,\Delta}{S(\infty)}\, d\mu,$$

а из последней по аналогии с (14.220) получим

$$h = \frac{1}{4\pi} \int_{-\infty}^{\infty} \ln \frac{S(\omega)}{S(\infty)}\, d\omega. \qquad (14.225)$$

Здесь подынтегральное выражение стремится к нулю при $|\omega|\to\infty$. Поскольку $\ln(1+y) = y - \frac{y^2}{2} + \ldots$, то интеграл в (14.225) сходится, если

$$\int_c^{\infty} \frac{S(\omega) - S(\infty)}{S(\infty)}\, d\omega < \infty, \qquad \int_c^{\infty} \left[\frac{S(\omega)}{S(\infty)} - 1\right]^2 d\omega < \infty \quad (14.226)$$

(c — некоторое число). Таким образом, мы получаем для удельной энтропии h конечное значение, если выполняется условие (14.226) и все особенности спектральной плотности (где она обращается в бесконечность или нуль) являются логарифмически интегрируемыми, как, например, нули и полюса вида

$$S(\omega) \sim |\omega - \omega_1|^r, \quad |r| < \infty.$$

При выполнении этих условий мера P является абсолютно непрерывной относительно специальным образом сконструированной меры v, определенной соотношением (14.223) и условием мультипликативности.

Условие сходимости другого интеграла (14.218) в верхнем пределе имеет вид

$$\int_c^{\infty} \left[\frac{S(\omega)}{\widetilde{S}(\omega)} - 1\right]^2 d\omega < \infty, \qquad (14.227)$$

аналогичный (14.216). Оно является необходимым условием абсолютной непрерывности меры P относительно Q.

Если для меры Q выполнено условие мультипликативности то (в стационарном случае) ее корреляционная матрица должна быть кратна единичной матрице, а спектральная плотность постоянна; $\tilde{S}(\omega) = \tilde{S}$. Для абсолютной непрерывности необходимо равенство $S(\infty)=\tilde{S}(\infty)$, следовательно, $\tilde{S}(\omega) = S(\infty)$. Подставляя это значение в (14.218), получаем при совпадающих средних формулу

$$h^{P/Q} = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} G\left(\frac{S(\omega)}{S(\infty)}\right) d\omega \quad (G(x) = (x - 1 - \ln x)/2). \quad (14.228)$$

Сходство последней с (14.225) является очевидным, различие заключается лишь в выборе подынтегральной функции.

Пример 1. Пусть имеется стационарный гауссов процесс со спектральной плотностью

$$S(\omega) = S_0 \frac{\omega^2 + \gamma^2}{\omega^2 + \beta^2}.$$

Вычислим для него удельные энтропии. Поскольку $S(\infty) = S_0$, то в соответствии с (14.225) имеем

$$h = \frac{1}{2\pi}\int\limits_0^{\infty} \ln\frac{\omega^2 + \gamma^2}{\omega^2 + \beta^2} d\omega = \frac{1}{2}(\gamma - \beta).$$

Применяя далее формулу (14.228), находим

$$h^{P/Q} = \frac{1}{2\pi}\int\limits_0^{\infty}\left(\frac{\omega^2 + \gamma^2}{\omega^2 + \beta^2} - 1\right) d\omega - \frac{1}{2\pi}\int\limits_0^{\infty} \ln\frac{\omega^2 + \gamma^2}{\omega^2 + \beta^2} =$$

$$= \frac{\gamma^2 - \beta^2}{4\beta} - h - \frac{1}{4\beta}(\gamma - \beta)^2. \quad (14.229)$$

Граничная энтропия Г, входящая в соотношения (14.206), (14.207), может быть вычислена методом, который предложен Стратоновичем. Для данного примера она оказывается такой:

$$\Gamma = \frac{1}{2}\ln\frac{(\gamma/\beta)^{1/2} + (\beta/\gamma)^{1/2}}{2}.$$

Пример 2. Пусть теперь рассматриваемый процесс имеет спектральную плотность $S(\omega) = S_0/(\omega^2 + \beta^2)$. Тогда применение формулы (14.225) не дает конечного результата. Чтобы получить конечную удельную энтропию по формуле (14.221), нужно подобрать подходящую спектральную плотность $\tilde{S}(\omega)$. Положим

$$\tilde{S}(\omega) = S_0/\omega^2. \quad (14.220)$$

Тогда интеграл в (14.221) сведется к интегралам, стоящим в (14.229) (при $\gamma = 0$), и мы будем иметь

$$h^{P/Q} = \beta/4. \qquad (14.231)$$

При выборе спектральной плотности (14.230) мера *Q,* описывающая случайный процесс η*(t),* не удовлетворяет условию мультипликативности. Однако это условие будет выполнено для производной $\dot{\eta}(t) = d\eta(t)/dt$, так как последняя будет иметь равномерную спектральную плотность S_0. При переходе к производной $\dot{\eta}(t)$ процесс ξ *(t)* также нужно заменить на его производную $\dot{\xi}(t)$. Можно считать поэтому, что результат (14.231) согласуется с условием мультипликативности.

3. Для стационарного гауссового процесса, как и для гауссовой последовательности, из линейного роста дисперсии $\mathbf{D}H(\xi_0^T)$ с ростом *Т* при условии линейного нарастания средней энтропии $H_{\xi_0^T}$ вытекает выполнение условия энтропийной устойчивости (14.18). Поэтому в случае ненулевой конечной удельной энтропии *h* для доказательства энтропийной устойчивости следует проверить конечность предела

$$D_0 = \lim_{T \to \infty} \frac{1}{T} \mathbf{D} H(\xi_0^T).$$

Чтобы его вычислить, нужно применить формулу (14.164). Диагонализуя входящие в нее матрицы, нетрудно получить (подобно тому, как из (14.160) было получено выражение (14.221)) следующее равенство:

$$D_0 = \frac{1}{2\pi} \int_0^\infty \left[\frac{S(\omega)}{\tilde{S}(\omega)} - 1 \right]^2 d\omega + \frac{(m - \tilde{m})^2 S(0)}{4\tilde{S}^2(0)}. \qquad (14.232)$$

Оно является обобщением на случай непрерывного времени равенства (14.189). Условие конечности интеграла в (14.232), очевидно, связано с указанным ранее условием (14.227). Таким образом, условие энтропийной устойчивости для гауссовых мер оказывается тесно связанным с условием абсолютной непрерывности меры *Р* относительно *Q*.

Вышеизложенное может быть обобщено и на случай нескольких стационарных и стационарно связанных процессов $\{\xi^\alpha(t)\}$, которые характеризуются столбцом средних значений $m = \|m^\alpha\|$ и матрицей спектральных плотностей

$$\mathrm{S}(\omega) = \|S^{\alpha\beta}(\omega)\|, \quad S^{\alpha\beta}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega\tau} \times$$

$$\times \mathbf{M}[\xi^\alpha(t) - m^\alpha][\xi^\beta(t+\tau) - m^\beta]\, d\tau.$$

Так, если мера Q описывается матрицами $\widetilde{\mathrm{m}}$, $\widetilde{\mathrm{S}}(\omega)$, то формула, обобщающая формулу (14.221), имеет вид

$$h^{P/Q}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{Sp} G\left(\widetilde{\mathrm{S}}^{-1}(\omega)\, S(\omega)\right) d\omega+$$

$$+\frac{1}{2}\left(\mathrm{m}^{T}-\widetilde{\mathrm{m}}^{T}\right)\widetilde{\mathrm{S}}(0)^{-1}\left(\mathrm{m}-\widetilde{\mathrm{m}}\right). \qquad (14.233)$$

Очевидна аналогия этой формулы с (14.187).

14.12. Энтропия точечного случайного процесса

Рассмотрим точечный случайный процесс на интервале $a \le t \le b$, представляющий собой совокупность случайных точек, положение и число которых случайно. Иногда подобный процесс называют «случайным потоком», однако, этот термин нам представляется неудачным, поскольку слово «поток» естественнее употреблять в другом смысле, связывая его с движением в пространстве, подобно тому, как это делается в физике.

Обозначим через n число выпавших точек, а через $\tau_1, \ldots, \tau_n$ места их выпадения (так что все $\tau_i \in [a, b]$). Будем предполагать, что величины $\tau_1, \ldots, \tau_n$ пронумерованы в порядке неубывания: $\tau_1 \leqslant \tau_2 \leqslant \ldots \leqslant \tau_n$. Совпадение $\tau_i = \tau_j$, $i \neq j$ предполагаем имеющим нулевую вероятность.

1. Для описанного процесса мера **P** характеризуется заданием вероятностей

$$\mathbf{P}(n) = P_n \qquad (14.234)$$

и системы плотностей распределения

$$p(\tau_1 \mid 1), \qquad p(\tau_1, \tau_2 \mid 2), \ldots, \qquad (14.235)$$

удовлетворяющих условию нормировки

$$\int_a^b d\tau_1 \int_{\tau_1}^b d\tau_2 \int_{\tau_2}^b d\tau_3 \ldots \int_{\tau_{n-1}}^b d\tau_n\, p(\tau_1, \ldots, \tau_n \mid n) = 1. \qquad (14.236)$$

Требуется определить энтропию данного точечного процесса. Вероятности (14.234) и плотности (14.236) определяют вероятность

$$\Delta P = P_n\, p_n(t_1, \ldots, t_n)\, \Delta_1 \ldots \Delta_n \left(1 + O(\Delta_{\max})\right) \qquad (t_1 < \ldots < t_n),$$

$$\Delta_{\max} = \max(\Delta_1, \ldots, \Delta_n) \qquad (14.237)$$

того, что первая точка τ_1 из n последовательных случайных точек $\tau_1 < \ldots < \tau_n$ попадет в интервал $[t_1, t_1 + \Delta_1]$, вторая — в интервал $[t_2, t_2 + \Delta_2]$ и т. д., последняя — в интервал $[t_n, t_n + \Delta_n]$

при условии, что в других местах интервала [*a, b*] не выпадет ни одной случайной точки. Простейшей системой случайных точек является пуассоновская система, для которой

$$P_n = \frac{1}{n!}\left(\int_a^b \beta(t)\,dt\right)^n \exp\left[-\int_a^b \beta(t)\,dt\right],$$

$$p_n(\tau_1, \ldots, \tau_n) = n!\,\beta(\tau_1)\ldots\beta(\tau_n) \Big/ \left[\int_a^b \beta(t)\,dt\right]^n. \quad (14.238)$$

Здесь $\beta\ (\tau)$ — средняя плотность точек, которая в стационарном случае постоянна.

Рассмотрение случайных точек $\tau_1, \ldots, \tau_n$ эквивалентно рассмотрению случайного процесса

$$\xi(t) = \sum_{j=1}^{n} \delta(t-\tau_j).$$

Выпадение случайных точек на неперекрывающихся интервалах $[\alpha, \beta], [\gamma, \delta]$ для пуассоновского процесса является взаимно независимым (или, иначе говоря, ξ_α^β независимы от ξ^δ_γ). Поэтому для выполнения условия мультипликативности (14.193) удобно брать энтропию точечного процесса, определенную в соответствии с (14.50), (14.52), выбирая меру *Q*, соответствующую пуассоновскому процессу. Будем полагать для простоты, что мере *Q* соответствует постоянная пуассоновская плотность β.

В соответствии с (14.238) для отношения элементарных вероятностей типа (14.237) в стационарном случае будем иметь

$$\frac{dP}{dQ} = \frac{P_n\,p_n(t_1, \ldots, t_n)}{\beta^n e^{-\beta(b-a)}}.$$

Поэтому энтропия (14.51) будет определяться формулой

$$H^{P/Q}_{\xi_a^b} = \mathbf{M}\ln\frac{dP}{dQ} = \sum_{n=0}^{\infty} \int_{a<t_1<\ldots<t_n<b} \ldots \int P_n\,p_n(t_1, \ldots, t_n) \times$$
$$\times \ln[P_n e^{\beta(b-a)} p_n(t_1, \ldots, t_n)\,\beta^{-n}]\,dt_1\ldots dt_n. \quad (14.239)$$

Выделяя β из-под знака логарифма, это выражение, очевидно, можно записать в виде

$$H^{P/Q}_{\xi_a^b} = (\beta-1)\,T - \ln\beta\,\mathbf{M}n + H^{P/Q_1}_{\xi_a^b} \quad (T=b-a), \quad (14.240)$$

$$H^{P/Q_1}_{\xi_a^b} = \sum_{n=0}^{\infty} \int_{a<t_1<\ldots<t_n<b} \ldots \int P_n\,p_n(t_1, \ldots, t_n) \times$$
$$\times \ln[e^T P_n\,p_n(t_1, \ldots, t_n)]\,dt_1\ldots dt_n. \quad (14.241)$$

— энтропия, соответствующая пуассоновской мере Q_1 с единичной плотностью.

Из энтропии (14.241) можно выделить энтропию H_n случайного числа точек

$$H_n = -\sum_{n=0}^{\infty} P_n \ln P_n, \qquad (14.242)$$

представив (14.241) в форме

$$H^{P/Q_1}_{\xi_a^b} = H^{P/Q_1}_{\xi_a^b|n} - H_n,$$

где

$$H^{P/Q_1}_{\xi_a^b|n} = \sum_{n=0}^{\infty} P_n \int_{a<t_1<\dots<t_n<b} \dots \int p_n(t_1, \dots, t_n) \times$$

$$\times \ln[e^T p_n(t_1, \dots, t_n)]\, dt_1 \dots dt_n. \qquad (14.243)$$

Вследствие общих свойств энтропии выражения (14.239), (14.241) — (14.243) всегда неотрицательны.

Возьмем для примера пуассоновскую систему точек с постоянной плотностью γ. Тогда в силу (14.238), (14.241) будем иметь

$$H^{P/Q_1}_{\xi_a^b} = \mathbf{M} \ln(e^{T-\gamma T} \gamma^n) = \gamma T \ln\gamma + (1-\gamma) T, \qquad (14.244)$$

поскольку $\mathbf{M}n = \gamma T$.

Далее из (14.240) получаем

$$H^{P/Q}_{\xi_a^b} = (\beta-\gamma)(b-a) + \gamma(b-a)\ln\frac{\gamma}{\beta}. \qquad (14.245)$$

Поделив этот результат на *b* — *a,* найдем удельную энтропию одной пуассоновской меры по другой:

$$h^{P/Q} = \beta - \gamma + \gamma\ln\frac{\gamma}{\beta} = \gamma\left(\frac{\beta}{\gamma} - 1 - \ln\frac{\beta}{\gamma}\right) \geqslant 0.$$

Энтропии (14.244), (14.245) пропорциональны *b* — *a* вследствие того, что пуассоновские процессы являются процессами с независимыми значениями. Более сложно зависят от длины интервала энтропии (14.242), (14.243). Вводя функцию

$$S(x) = e^{-x} \sum_{n=0}^{\infty} \frac{x^n}{n!} \ln n!,$$

эти выражения для пуассоновской меры *P* можно записать

$$H_n = \gamma T[1 - \ln(\gamma T)] + S(\gamma T), \qquad H^{P/Q_1}_{\xi_0^T|n} = T + S(\gamma T) - \gamma T \ln T.$$

2. Приведенные выше формулы основывались на определении (14.51) энтропии одной вероятностной меры по другой. Можно перейти к энтропии (14.47), (14.50), соответствующей ненормированной мере *v*. Эта мера будет удовлетворять условию мультинликативности, если будет пропорциональна пуассоновской вероятностной мере *Q*.

Согласно сказанному ранее обобщенную энтропию (14.50) можно интерпретировать как предельный случай энтропии дискретной версии, в которой число возможных исходов является конечным или счетным. Чтобы перейти от случая непрерывного времени *t* к дискретному случаю, разобьем рассматриваемый интервал на конечное число элементарных интервалов $[t'_m,\ t'_{m+1})$ и вместо интервала [*a,* *b*] будем рассматривать множество Z точек $t'_0 = a;\ t'_1,\ \ldots,\ t'_{N-1};\ t'_N = b$. Точечный процесс на Z определим так:

пусть $\xi'(t'_m) = 1$, если на интервал $[t'_m, t'_{m+1})$ попадает хотя бы одна точка τ_k и $\xi'(t'_m) = 0$ в противном случае. Имея в виду рассмотреть в дальнейшем стационарный процесс, будем полагать разбиение равномерным:

$$t'_{m+1} - t'_m = \Delta, \qquad m = 0,\ 1,\ \ldots,\ N-1 \qquad (N = (b-a)/\Delta).$$

Вероятность выпадения на элементарном интервале больше одной точки предполагается величиной порядка Δ^2, т. е. высшего порядка малости. Тогда при достаточно большом *N* число *n* случайных точек, выпавших в [*a,* *b*] с вероятностью, близкой к единице, не будет отличаться от числа точек

$$n' = \sum_{k=0}^{N-1} \xi'(t'_k),$$

выпавших в *Z*.

Точечный процесс на Z является дискретным процессом, имеющим в общей сложности 2^N различных реализаций. Событие $n'=0$ реализуется одним способом. Событие $n'=1$ реализуется *N* различными способами, т. е. включает *N* различных реализаций. Событие $n' = 2$ состоит из $N(N - 1)/2$ различных реализаций и т. д. Вводя меру, указывающую число различных реализаций, имеем

$$\nu(n'=0) = 1, \qquad \nu(n'=1) = N, \quad \nu(n'=2) = N(N-1)/2!\ , \ldots$$
$$\ldots,\ \nu(n') = N!/n'!\,(N-n')!\ ,\ \ldots$$

Причем

$$\sum_{n'=0}^{\infty} \nu(n') = 2^N$$

Нетрудно подсчитать также число реализаций того события, что τ_1 попадет в интервал $[t_1, \ t_1 + \Delta_1)$, в то же время τ_2 попадает в интервал $[t_2, \ t_2 + \Delta_2)$ и т. д., τ_n попадет в интервал $[t_n, \ t_n + \Delta_n)$.

В предположении, что $\Delta_k \gg \Delta, \ n \ll N$, число таких реализаций равно

$$\Delta\nu \approx \frac{\Delta_1}{\Delta} \ldots \frac{\Delta_n}{\Delta}.$$

Напомним, что вероятность такого множества задается формулой (14.237). Поскольку

$$\frac{\Delta P}{\Delta\nu} \approx P_n \, p_n(t_1, \ldots, t_n)\, \Delta^n,$$

то применяя формулы (14.39), (14.40), получаем

$$H_{\xi'} \approx -\sum P_n \, p_n(t_1, \ldots, t_n) \ln [P_n \, p_n(t_1, \ldots, t_n)\, \Delta^n].$$

Выноса Δ из-под знака логарифма, полученный результат можно выразить через энтропию (14.241):

$$H_{\xi'} \approx \ln \frac{1}{\Delta} \mathbf{M}n + b - a - H_{\xi_a^b}^{P/Q_1}. \qquad (14.246)$$

Это обосновывает введение энтропии H_{ξ}^{P/Q_1} меры P относительна пуассоновской меры Q_t. Если разбиение интервала $[a, \ b]$ производить не равномерным образом, а согласно соотношению

$$t_{m+1} - t_m = \beta\,(t_m)\Delta,$$

то энтропия $H_{\xi'}$ по аналогии с (14.246) будет выражаться через энтропию $H_{\xi}^{P/Q}$ меры P по пуассоновской мере Q, имеющей неравномерную плотность $\beta(t)$.

3. В случае стационарного точечного процесса можно рассматривать удельную энтропию, приходящуюся в среднем на единицу времени. Укажем два способа ее вычисления.

1) В соответствии с теоремой 12 ее можно вычислять по формуле (14.200). Имея в виду применить эту формулу, исследуем поведение энтропии (14.241), (14.243) при больших n.

В случае стационарного процесса среднее число точек $\mathbf{M}n \equiv \bar{n}$ пропорционально длине интервала $b - a = T$. Для эргодического процесса, кроме того, зависимость дисперсии $\mathbf{D}n$ от T при больших n, как правило, приближается к линейной:

$$\mathbf{D}n = D_0 T + O(1),$$

и случайная величина n подчиняется центральной предельной теореме. Вероятность неравенства $n_0 \leqslant n \leqslant n_0 + \Delta N$ при больших n может быть вычислена при помощи гауссового распределения:

$$\Delta P \approx \frac{1}{\sqrt{2\pi \mathbf{D}n}} e^{-(n_0 - \bar{n})^2/2\mathbf{D}n} \Delta N \qquad (\Delta N^2 \ll \mathbf{D}n), \quad (14.247)$$

где ΔN имеет тот же смысл, что и в (14.38). Это позволяет вычислить H_n приближенно как энтропию гауссовой переменной. Из (14.247) имеем

$$H(n) = -\ln P_n \approx -\ln \frac{\Delta P}{\Delta N} \approx \frac{1}{2} \ln (2\pi \mathbf{D}n) + \frac{(n - \bar{n})^2}{2\mathbf{D}n}, \quad (14.248)$$

что является частным случаем формулы (14.150). Усредняя (14.248), находим

$$H_n \approx \frac{1}{2} \ln 2\pi e\, \mathbf{D}n \approx \frac{1}{2} \ln (2\pi e D_0 T).$$

Из найденной зависимости следует, в частности, исчезновение предела

$$\lim_{t \to \infty} \frac{1}{T} H_n = 0.$$

Это означает, что вклад энтропии H_n в удельную энтропию (14.200) сводится к нулю. На удельную энтропию, следовательно, оказывает влияние лишь энтропия (14.243):

$$h^{P/Q_1} = \lim_{T \to \infty} \frac{1}{T} H^{P/Q_1}_{\xi_0^T | n}. \quad (14.249)$$

Вводя обозначение

$$\Phi(n) = H^{P/Q_1}_{\xi_0^T}(| n) = \int_{0 \leqslant t_1 <} \cdots \int_{< t_n < T} p_n(t_1, \ldots, t_n) \times$$
$$\times \ln [e^T p_n(t_1, \ldots, t_n)]\, dt_1 \ldots dt_n, \quad (14.250)$$

ее можно записать так

$$H^{P/Q_1}_{\xi_0^T | n} = \mathbf{M}\Phi(n). \quad (14.251)$$

Величина n при больших T является квазинепрерывной. Для нее n мало отличается от $[\bar{n}]$ (скобки [...] означают целую часть). Разности

$$\Phi'(n) = \Phi(n+1) - \Phi(n),$$
$$\Phi''(n) = \Phi(n+1) - 2\Phi(n) + \Phi(n-1)$$

можно интерпретировать как производные. Пренебрегая несущественными усложнениями, связанными с дискретным характером величины *n,* из (14.250) будем иметь

$$\mathbf{M}\Phi(n) = \mathbf{M}\left[\Phi(\bar{n}) + \Phi'(\bar{n})(n-\bar{n}) + \frac{1}{2}\Phi''(\bar{n})(n-\bar{n})^2 + \ldots\right] =$$

$$= \Phi(\bar{n}) + \frac{1}{2}\Phi''(\bar{n})\mathbf{D}n + \ldots .$$

Подставляя это выражение в (14.151) и (14.249), получаем, что

$$h^{P/Q_t} = \lim_{T\to\infty} \frac{1}{T}\Phi(\bar{n}), \qquad (14.252)$$

если

$$\lim_{T\to\infty} \Phi''(\bar{n})\frac{\mathbf{D}n}{T} = 0. \qquad (14.253)$$

Уже указывалось, что обычно отношение $\mathbf{D}n/T$ стремится при $T\to\infty$ к конечному пределу D_0. Следовательно, условие (14.253) выполняется, если

$$\Phi''(\bar{n}) \to 0 \text{ при } T\to\infty. \qquad (14.254)$$

Как показывает исследование, условие (14.254), говорящее о приближении зависимости $\Phi(n)$ при $n = cT$ к прямой пропорциональности, выполняется для большого числа практических случаев.

Итак, в соответствии с формулой (14.252) удельную энтропию можно вычислить, считая число случайных точек, выпавших на большом интервале $[0, T]$, неслучайным, заранее известным и равным $\bar{n} = \mathbf{M}n$.

Пример. Вычислим этим способом энтропию пуассоновского процесса. Предполагается, что число точек на всем интервале является неслучайным и равным $\bar{n} = \gamma T$. При этом $p_n(t_1 \ldots, t_n) = n!T^{-n}$. Подставляя эту функцию в (14.250), находим

$$\Phi(n) = T + \ln n! - n\ln T.$$

Пользуясь формулой Стирлинга

$$n! = n^n e^{-n}\sqrt{2\pi n}\,[1 + O_n(1)],$$

получаем

$$\Phi(n) = T + n\ln n - n + \frac{1}{2}\ln(2\pi n) - n\ln T.$$

Условие (14.254) действительно имеет место, поскольку

$$\Phi''(\bar{n}) \approx \frac{1}{\bar{n}} = \frac{1}{\gamma T}.$$

Подставляя найденное выражение для Ф (n) в (14.252) при $\bar{n} = \gamma T$ и перехояя к пределу $T\to\infty$, имеем

$$h^{P/Q_1} = 1 + \gamma \ln\gamma - \gamma, \tag{14.255}$$

что совпадает с удельной энтропией, находимой из (14.244).

2) Другой способ вычисления удельной энтропии основан на ее определении (14.199) как условной энтропии. Фиксация процесса $\xi_{-\infty}^{0}$ означает фиксацию всех случайных точек $\ldots, \tau_{-2}, \tau_{-1}, \tau_0 < 0$, выпавших до момента $t = 0$. Поэтому (14.199) можно записать:

$$h^{P/Q_1} = \frac{1}{\tau} H^{P/Q_1}_{\xi_0^\tau \mid \ldots, \tau_{-1}, \tau_0} = \frac{1}{\tau} M H^{P/Q_1}_{\xi_0^\tau} (\mid \ldots, \tau_{-1}, \tau_0). \tag{14.256}$$

Энтропию $H^{P/Q}_{\xi_0^\tau} (\mid \ldots, \tau_{-1}, \tau_0)$, здесь следует определять формулой (14.241), но при замене вероятностей P_n и плотностей $p_n (\tau_1, \ldots, \tau_n)$ условными вероятностями и плотностями

$$P(n \mid \ldots, \tau_{-1}, \tau_0), \; p_n (\tau_1, \ldots, \tau_n \mid \ldots, \tau_{-1}, \tau_0), \tag{14.257}$$

т.е. формулой

$$H^{P/Q_1}_{\xi_0^\tau} (\mid \ldots, \tau_{-1}, \tau_0) = \sum_{n=0}^{\infty} P(n \mid \ldots, \tau_{-1}, \tau_0) \times$$
$$\times \int_{t_1 <} \ldots \int_{t_n} p_n (t_1, \ldots, t_n \mid \ldots, \tau_1, \tau_0) \ln [e^\tau P_n (n \mid \ldots, \tau_{-1}, \tau_0) \times$$
$$\times p_n (t_1, \ldots, t_n \mid \ldots, \tau_{-1}, \tau_0)] \, dt_1 \ldots dt_n. \tag{14.258}$$

Длину τ интервала [0, τ] желательно брать небольшой, чтобы вероятности выпадения двух и более точек на этом интервале были пренебрежимо малы и их можно было бы не учитывать. Тогда в выражении (14.257) можно будет оставить лишь два члена, полагая при этом $p_1 (t_1 \mid \ldots, \tau_{-1}, \tau_0) = 1/\tau$. Обозначая кроме того

$$\lim_{\tau\to 0} \mathbf{P}(n = 1 \mid \ldots, \tau_{-1}, \tau_0)/\tau = p(0 \mid \ldots, \tau_{-1}, \tau_0),$$

из (14.257) будем иметь

$$H^{P/Q_1}_{\xi_0^\tau} (\mid \ldots, \tau_{-1}, \tau_0) - [1 - p(0 \mid \ldots, \tau_{-1}, \tau_0)\tau] \times$$
$$\times \ln \{e^\tau [1 - p(0 \mid \ldots, \tau_{-1}, \tau_0)\tau]\} +$$
$$+ p(0 \mid \ldots, \tau_{-1}, \tau_0)\tau \ln [e^\tau p(0 \mid \ldots, \tau_{-1}, \tau_0)] + O(\tau^2),$$

или

$$H^{P/Q_1}_{\xi_0^\tau} (\mid \ldots, \tau_{-1}, \tau_0) = \tau - p(0 \mid \ldots, \tau_{-1}, \tau_0)\tau +$$
$$+ p(0 \mid \ldots, \tau_{-1}, \tau_0)\tau \ln p(0 \mid \ldots, \tau_{-1}, \tau_0) + O(\tau^2). \tag{14.259}$$

Подставляя это выражение в (14.256) и переходя к пределу $\tau \to 0$, получаем искомую удельную энтропию

$$h^{P/Q_1} = \mathbf{M}[1 - p(0|\ldots, \tau_{-1}, \tau_0) + p(0|\ldots, \tau_{-1}, \tau_0) \times \times \ln p(0|\ldots, \tau_{-1}, \tau_0)], \qquad (14.260)$$

где **M** обозначает усреднение по случайным точкам $\ldots, \tau_{-1}, \tau_0$.

Пример 1. Применим найденную формулу к стационарному точечному процессу с ограниченным последействием. Будем полагать, что интервалы $\sigma = t_{m+1} - t_m$ между соседними случайными точками являются независимыми случайными величинами, имеющими одинаковую плотность распределения вероятностей $w(\sigma)$. Тогда, очевидно,

$$\mathbf{P}(n = 1 | \ldots, \tau_{-1}, \tau_0) = w(-\tau_0)\tau \Big/ \int_{-\tau_0}^{\infty} w(\sigma)\, d\sigma + O(\tau^2),$$

так что

$$p(0|\ldots, \tau_{-1}, \tau_0) = w(-\tau_0) \Big/ \int_{-\tau_0}^{\infty} w(\sigma)\, d\sigma.$$

Подстановка в (14.260) дает

$$h^{P/Q_1} = 1 + \int_0^{\infty} \left[\ln \frac{w(\sigma)}{\int_{\sigma}^{\infty} w(\rho)\, d\rho} - 1 \right] \frac{w(\sigma)}{\int_{\sigma}^{\infty} w(\rho)\, d\rho} P(d\sigma), \qquad (14.261)$$

где $P(d\sigma) = \mathbf{P}[\sigma \leqslant -\tau_0 < \sigma + d\sigma] \cong p(\sigma)\, d\sigma$ — закон распределения для случайного расстояния от фиксированной точки $t = 0$ до ближайшей слева случайной точки. Его можно выразить через плотность $w(\sigma)$ формулой

$$p(\sigma)\, d\sigma = \frac{d\sigma}{\bar{\sigma}} \int_{\sigma}^{\infty} w(\rho)\, d\rho, \qquad \bar{\sigma} = \int \rho w(\rho)\, d\rho. \qquad (14.262)$$

Вследствие равенств (14.261), (14.262) получаем

$$h^{P/Q_1} = 1 + \frac{1}{\bar{\sigma}} \int_0^{\infty} \left[\ln w(\sigma) - \ln \int_{\sigma}^{\infty} w(\rho)\, d\rho - 1 \right] w(\sigma)\, d\sigma, \qquad (14.263)$$

что дает решение задачи вычисления удельной энтропии.

Пример 2. Рассмотрим несколько более сложный пример. Пусть дан стационарный точечный процесс. Интервалы $\tau_{m+1} - \tau_m$, по-прежнему, взаимно независимы, но имеют (через один) различные плотности распределения; $w_1(\sigma)$ или $w_2(\sigma)$. Если $\tau_{m+1} - \tau_m$ распределен по w_1, то $\tau_{m+2} - \tau_{m+1}$ распределен по w_2; интервал

$\tau_{m+3} - \tau_{m+2}$ снова распределен по w_1, а $\tau_{m+4} - \tau_{m+3}$ — по w_2 и т. д. Такой точечный процесс эквивалентен стационарному процессу с двумя состояниями A_1 и A_2, когда времена пребывания в каждом состоянии взаимно независимы и имеют законы распределения w_1 (для времени пребывания в A_1) и w_2 (в A_2). Случайные точки можно при этом классифицировать дополнительным параметром ϑ, полагая $\vartheta_m = 1$, если в точке τ_m, происходит скачок из A_1 в A_2, и $\vartheta_m = 2$, если происходит обратный скачок из A_2 в A_1.

В описанном случае плотность вероятности $p(0 \mid \ldots, \tau_{-1}, \tau_0)$ зависит не только от времени τ_0 выпадения последней случайной точки, но и от ее типа ϑ_0. Именно

$$p(0 \mid \ldots, \tau_{-1}, \tau_0) = p(0 \mid \tau_0, \vartheta_0) = w_{\vartheta_0}(-\tau_0) \Big/ \int\limits_{-\tau_0}^{\infty} w_{\vartheta_0}(\rho)\, d\rho. \tag{14.264}$$

Усреднение в (14.260) будет проводиться как по τ_0, так и по ϑ_0. Обозначим через $P(\vartheta, d\sigma)$ совместное распределение для случайных величин $\vartheta = \vartheta_0$, $\sigma = -\tau_0$. Тогда из (14.260), (14.264) будем иметь

$$h^{P/Q_1} = 1 + \sum_{\vartheta=1}^{2} \int \left[\ln \frac{w_\vartheta(\sigma)}{\int\limits_\sigma^\infty w_\vartheta(\rho)\, d\rho} - 1 \right] \times$$

$$\times \frac{w_\vartheta(\sigma)}{\int\limits_\sigma^\infty w_\vartheta(\rho)\, d\rho} P(\vartheta, d\sigma). \tag{14.265}$$

Остается вычислить $P(\vartheta, d\sigma)$. Априори вероятность попадания на элементарный интервал $[-\sigma, -\sigma + d\sigma]$ случайной точки любого из двух типов одна и та же и равна

$$dP = \frac{d\sigma}{\bar\sigma_1 + \bar\sigma_2}, \qquad \bar\sigma_\vartheta = \int\limits_0^\infty \rho w_\vartheta(\rho)\, d\rho,$$

поскольку для каждого типа средняя плотность точек равна $(\bar\sigma_1 + \bar\sigma_2)^{-1}$. Если выпала точка $\vartheta_0 = 1$, то с вероятностью

$$\int\limits_\sigma^\infty w_1(\rho)\, d\rho$$

она будет последней, т. е. в $[-\sigma + d\sigma, 0]$ не выпадет ни одной другой точки. Если $\vartheta_0 = 2$, то точка τ_0 будет последней с вероятностью $\int\limits_{\sigma}^{\infty} w_2(\rho)\, d\rho$. Поэтому формулу (14.265) можно записать

$$h^{P/Q_1} = 1 + \frac{1}{\bar{\sigma}_1 + \bar{\sigma}_2} \sum_{\vartheta=1}^{2} \int\limits_{0}^{\infty} \left[\ln w_\vartheta(\sigma) - \ln \int\limits_{\sigma}^{\infty} w_\vartheta(\rho)\, d\rho - 1 \right] \times$$
$$\times\, w_\vartheta(\sigma)\, d\sigma. \qquad (14.266)$$

Пусть, например w_1, w_2 имеют экспоненциальный вид

$$w_\vartheta(\sigma) = \mu_\vartheta\, e^{-\mu_\vartheta \sigma}.$$

Тогда

$$\bar{\sigma}_\vartheta = \frac{1}{\mu_\vartheta}, \qquad \int\limits_{\sigma}^{\infty} w_\vartheta(\rho)\, d\rho = e^{-\mu_\vartheta \sigma},$$

и из (14.266) получаем

$$h^{P/Q_1} = 1 + \frac{\mu_1 \mu_2}{\mu_1 + \mu_2} \sum_{\vartheta=1}^{2} [\ln \mu_\vartheta - 1] =$$
$$= 1 - \frac{2\mu_1 \mu_2}{\mu_1 + \mu_2} + \frac{\mu_1 \mu_2}{\mu_1 + \mu_2} (\ln \mu_1 + \ln \mu_2). \qquad (14.267)$$

Если, в частности, $\mu_1 = \mu_2 = \gamma$, то формула (14.267) совпадает с (14.255), так как при этом рассматриваемый точечный процесс переходит в пуассоновский.

14.13. Энтропия дискретного марковского процесса в непрерывном времени

Рассмотрим марковский процесс $\xi(t)$ с дискретным пространством состояний, т. е. имеющий конечное или счетное число возможных состояний. В отличие от процесса, рассмотренного ранее, он протекает теперь в непрерывном времени. Пусть его вероятности переходов определяются дифференциальной вероятностью перехода $\pi_t(x, x')$ в соответствии с формулой

$$P[\xi(t+\Delta) \mid \xi(t)] = \delta_{\xi(t+\Delta),\, \xi(t)} + \pi_t(\xi(t), \xi(t+\Delta))\, \Delta + O(\Delta^2).$$

В силу условия нормировки имеем

$$\sum_{x'} \pi_t(x, x') = 0,$$

что соответствует (14.82).

Если процесс $\xi(t)$ стационарный, то матрица $\pi_t(x, x') = \pi(x, x')$ не зависит от времени t.

На данный марковский процесс могут быть распространены методы и результаты, изложенные ранее.

Любая конечная совокупность случайных значений $\xi(t_1), \ldots, \xi(t_k)$ является для данного процесса совокупностью дискретных случайных величин. Поэтому их энтропию можно вычислить по формулам дискретной версии. В частности, если процесс задан на интервале $[0, T]$, энтропия начального значения $\xi(0)$ равна

$$H_{\xi(0)} = -\sum_{\xi(0)} P(\xi(0)) \ln P(\xi(0)). \tag{14.268}$$

Если, однако, мы хотим вычислить энтропию непрерывного множества значений ξ^T_0, то мы должны применить формулы обобщенной версии.

Обозначим через $\{\tau_j\}$ те точки временной оси, т. е. те моменты времени, в которых происходят скачкообразные изменения состояния процесса. Вспомогательную меру $Q_1(\xi^T_0)$ определим как пуассоновскую меру (с единичной плотностью) для системы случайных точек $\{\tau_j\}$. Выделяя энтропию начального значения (14.268) по формуле

$$H_{\xi_0^T} = H_{\xi(0)} + H_{\xi_{0+0}^T \mid \xi(0)},$$

условную энтропию $H_{\xi_{0+0}^T \mid \xi(0)}$ или короче

$$H_{\xi_0^T \mid \xi(0)}$$

определим формулами (14.190) обобщенной версии. Тогда будем иметь

$$H^{P/Q_1}_{\xi_0^T} = -H_{\xi(0)} + H^{P/Q_1}_{\xi_0^T \mid \xi(0)}. \tag{14.269}$$

Вследствие мультипликативного свойства меры Q_1 энтропия H^{P/Q_1} обладает свойством иерархической аддитивности

$$H^{P/Q_1}_{\xi_\alpha^\gamma \mid \xi(a)} = H^{P/Q_1}_{\xi_\alpha^\beta \mid \xi(a)} + H^{P/Q_1}_{\xi_\beta^\gamma \mid \xi_\alpha^\beta}.$$

Согласно условию марковости процесса последнюю формулу можно записать

$$H^{P/Q_1}_{\xi^\gamma_\alpha | \xi(\alpha)} = H^{P/Q_1}_{\xi^\beta_\alpha | \xi(\alpha)} + H^{P/Q_1}_{\xi^\gamma_\beta | \xi(\beta)}.$$

Так что

$$H^{P/Q_1}_{\xi^T_0} = -H_{\xi(0)} + H^{P/Q_1}_{\xi^{t_1}_0 | \xi(0)} + H^{P/Q_1}_{\xi^{t_2}_{t_1} | \xi(t_1)} + \ldots +$$

$$+ H^{P/Q_1}_{\xi^T_{t_N} | \xi(t_N)} \qquad (0 < t_1 < \ldots < t_N < T). \qquad (14.270)$$

Это равенство является аналогом равенства (14.86), относящегося к случаю дискретного времени.

Как отмечалось ранее, для стационарного процесса $\xi\,(t)$ энтропия

$$H^{P/Q_1}_{\xi^\tau_0 | \xi^0_{-\infty}}$$

пропорциональна τ. Вследствие условия Маркова это относится и к энтропии $H^{P/Q_1}_{\xi^\tau_0 | \xi(0)}$. Формула (14.199) принимает вид

$$H^{P/Q_1}_{\xi^\tau_0 | \xi(0)} = \tau h^{P/Q_1}. \qquad (14.271)$$

Учитывая это, из (14.269) или (14.271) имеем

$$H^{P/Q_1}_{\xi^T_0} = -H_{\xi(0)} + h^{P/Q_1} T.$$

Сравнивая эту формулу с (14.207), видим, что в данном случае $o_t(1) = 0$ и

$$2\Gamma = H_{\xi(0)} = -\sum_{\xi(0)} P(\xi(0)) \ln P(\xi(0)). \qquad (14.272)$$

Соотношением (14.271) удобно воспользоваться при малых τ, чтобы вычислить удельную энтропию h^{P/Q_1} по аналогии с тем, как это было сделано в 14.12 (см. (14.256), (14.258)).

При малых τ можно пренебречь вероятностью выпадения на (0, τ] больше чем одной точки перехода. Тогда останутся лишь такие возможности: не произойдет перехода с вероятностью

$$1 + \pi(\xi(0), \xi(0))\tau + O(\tau^2)$$

или произойдет переход в состояние $x' \neq \xi(0)$ с вероятностью $\pi(\xi(0), x')\,\tau + O(\tau^2)$. Аналогично (14.259) записываем энтропию этих событий:

$$H^{P/Q_1}_{\xi_0^\tau}\left(|\xi(0)\right) =$$
$$= \left[1+\pi\left(\xi(0),\ \xi(0)\right)\tau\right]\ln\left\{e^\tau\left[1+\pi\left(\xi(0),\ \xi(0)\right)\tau\right]\right\}+$$
$$+\sum_{\xi'\neq\xi(0)}\pi\left(\xi(0),\ \xi'\right)\tau\ln\left[e^\tau\pi\left(\xi(0),\ \xi'\right)\right]+O(\tau^2).$$

Усредняя по ξ(0) и совершая предельный переход

$$h^{P/Q_1} = \lim_{\tau\to 0}\frac{1}{\tau}H^{P/Q_1}_{\xi_0^\tau|\xi(0)},$$

отсюда получаем

$$h^{P/Q_1} = 1+\sum_\xi\left[\pi(\xi,\ \xi)+\sum_{\xi'\neq\xi}\pi(\xi,\ \xi')\ln\pi(\xi,\ \xi')\right]P(\xi). \quad (14.273)$$

Учитывая, что $\pi(\xi,\ \xi) = -\sum_{\xi'\neq\xi}\pi(\xi,\ \xi')$, в силу (14.273) имеем

$$h^{P/Q_1} = \sum_\xi P(\xi)\left(1+\sum_{\xi'\neq\xi}\left[\pi(\xi,\ \xi')\ln\pi(\xi,\ \xi')-\pi(\xi,\ \xi')\right]\right).$$

Здесь, как и в (14.173), $P(\xi)$ — стационарные вероятности, определяемые из уравнения

$$\frac{dP(\xi')}{dt} = \sum_\xi P(\xi)\,\pi(\xi,\ \xi') = 0. \quad (14.274)$$

Последнее уравнение аналогично (14.87), а формула (14.273) является обобщением на случай непрерывного времени приведенной ранее формулы (14.88).

Пример. Пусть имеется процесс с двумя состояниями, характеризуемый дифференциальными вероятностями перехода

$$\begin{pmatrix}\pi(1,1) & \pi(1,2)\\ \pi(2,1) & \pi(2,2)\end{pmatrix} = \begin{pmatrix}-\mu & \mu\\ \nu & -\nu\end{pmatrix}. \quad (14.275)$$

Уравнение (14.274) принимает вид

$$-\mu P(1)+\nu P(2) = 0,$$

так что стационарные вероятности оказываются такими:

$$P(1) = \frac{\nu}{\mu+\nu}, \qquad P(2) = \frac{\mu}{\mu+\nu} \quad (14.276)$$

(ср. с (14.99)). По формуле (14.273) записываем удельную энтропию

$$h^{P/Q_1} = 1+\frac{\mu\nu}{\mu+\nu}\{\ln\mu+\ln\nu-2\}.$$

Этот результат совпадает с (14.267) и это естественно, поскольку рассмотренный в примере 2 п.14.12 процесс становится марковским процессом с двумя состояниями, если распределения

$w_1(\sigma)$, $w_2(\sigma)$ экспоненциальные. При других видах распределений процесс с двумя состояниями будет немарковским, но станет марковским, если к $\xi(t)$ добавить еще одну переменную, а именно время $\tau=t-\tau_j$, прошедшее после момента τ_j последнего перескока. Комбинированный процесс $\{\xi(t), \sigma(t)\}$ будет марковским. Пространством состояний для него будут две полупрямых. Несмотря на такое усложнение, к нему применима почти без усложнений описанная выше теория.

Сформулируем в заключение этого пункта указанное обобщение на произвольное пространство в более полном виде.

Пусть пространство состояний X марковского процесса произвольно. Дифференциальная вероятность перехода $\pi(x, x')$ такова, что из каждой точки $x \in X$ возможны переходы лишь в конечное или счетное множество точек. Тогда удельная энтропия определяется формулой

$$h^{P/Q}=\int P_{\text{ст}}(d\xi)\left[1+\sum_{\xi'\neq\xi}\pi(\xi, \xi')(\ln\pi(\xi, \xi')-1)\right], \quad (14.277)$$

где $P_{\text{ст}}(d\xi)$— стационарное распределение, определяемое из уравнения

$$\int P(d\xi)\sum_{\xi'\in A}\pi(\xi, \xi')=0$$

(A произвольно), обобщающего (14.274). В данном случае формула (14.268) может оказаться несправедливой, и для определения $H_{\xi(0)}$ может потребоваться мера Q.

Приведенные формулы (14.273), (14.277) можно применять не только в стационарном случае для вычисления средней по времени удельной энтропии. Они пригодны также вследствие марковских свойств процесса и в нестационарном случае, а именно, определяют плотность энтропии $h(t)$, рассчитанную на единицу времени, которая может зависеть от времени t. При этом усреднение в них следует производить с нестационарным распределением $P(d\xi)$.

14.14. Энтропия диффузионных марковских процессов

Пусть $\{x, (t), t\in[0, T]\}$— соответствующий мере P диффузионный марковский процесс, характеризуемый сносом $a(x, t)$ и локальной дисперсией $b(x, t) > 0$. Он описывается уравнением Фоккера — Планка

$$\dot{p}_t(x) = -\frac{\partial}{\partial x}[ap_t(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}[bp_t(x)], \qquad (14.278)$$

где

$$p_t(x)\,dx = \mathbf{P}\{x_t \in [x, x + dx]\}.$$

Чтобы определить для него энтропию $H_{x_0^T}$, нужно подобрать меру Q, относительно которой мера P является абсолютно непрерывной. При этом желательно, чтобы она была максимально простой. Известно, что такой мерой является мера, соответствующая диффузионному процессу с той же локальной дисперсией *b (x, t)* и нулевым сносом, т. е. мера, для которой (14.278) заменяется уравнением

$$\dot{q}_t(x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}[b_t q(x)] \quad (q_t(x)\,dx = \mathbf{Q}\{x_t \in [x, x + dx]\}).$$

При этом производная Радона — Никодима имеет вид

$$\frac{P(dx_0^T)}{Q(dx_0^T)} = \frac{p(x(0))}{q(x(0))}\exp\left\{\int_0^T \frac{a(x(t),\ t)}{b(x(t),\ t)}\left[d^*x(t) - \right.\right.$$

$$\left.\left. -\frac{1}{2}a\left(x(t),\ t\right)dt\right]\right\}, \qquad (14.279)$$

где стохастический интеграл $\int \ldots d^*x(t)$ понимается в смысле Ито.

Постараемся теперь удовлетворить условию мультипликативности (14.193), чтобы была применима теория п. 14.10. Для этого представим процесс $\{x(t),\ t \in [0,\ T]\}$ как процесс $\{\xi(0),\ \xi(t),\ t \in (0,\ T]\}$, где $\xi(0) = x(0),\ \xi(t) = x(t),\ \ t > 0$, и потребуем, чтобы $b(x, t) = b(t)$ не зависела от x. Тогда мера Q будет соответствовать гауссовому дельта-коррелированному процессу:

$$\int \xi(t)\,dQ = 0, \qquad \int \xi(t)\,\xi(t')\,dQ = b(t)\,\delta(t - t'),$$

так что условие мультипликативности (14.193) будет выполнено. Используя (14.279), нетрудно вычислить энтропию (14.190), а также в предположении стационарности удельную энтропию

$$h^{P/Q} = \frac{1}{\tau} H^{P/Q}_{\xi_t^{t+\tau} \mid \xi_0^t} = \frac{1}{\tau} H^{P/Q}_{x_t^{t+\tau} \mid x(t)}.$$

При этом энтропию

$$H^{P/Q}_{x_t^{t+\tau} \mid x(t)}$$

удобно вычислять приближенно, пользуясь малостью τ, а затем совершить предельный переход

$$h^{P/Q} = \lim_{\tau\to 0} \frac{1}{\tau} H^{P/Q}_{x_t^{t+\tau}}\Big|_{x(t)}. \qquad (14.280)$$

При малых τ из (14.279) имеем

$$\mathbf{M} \ln \frac{P\left[dx_t^{t+\tau} \mid x(t)\right]}{Q\{d\,[x(\sigma) - x(t)],\ \sigma \in (t,\ t+\tau]\}} = \mathbf{M} \int_t^{t+\tau} \frac{a(x(\sigma))}{b} \Big[d^* x(\sigma) -$$

$$- \frac{1}{2} a\,\big(x(\sigma)\big)\, d\sigma\Big] \approx \mathbf{M} \frac{a\,(x\,(t))}{b} \Big[x(t+\tau) - x(t) -$$

$$- \frac{1}{2} a\,\big(x(t)\big)\, \tau\Big] + o(\tau).$$

Для вычисления $H^{P/Q_1}_{x_t^{t+\tau}\,|x(t)}$ проводим усреднение

$$H^{P/Q}_{x_t^{t+\tau}}\Big|_{x(t)} = \frac{a\,(x\,(t))}{b} \Big\{\mathbf{M}\Big[x(t+\tau) - x(t)\Big] -$$

$$- \frac{1}{2} a\,\big(x(t)\big)\,\tau\Big\} + o(\tau).$$

Учитывая, что в соответствии с определением сноса *a* справедливо соотношение

$$\mathbf{M}\ [x(t+\tau) - x(t) \mid x(t)] = a(x(t))\,\tau + o\ (\tau),$$

будем иметь

$$\frac{1}{\tau} H^{P/Q}_{x_t^{t+\tau}}\Big|_{x(t)} = \frac{1}{2}\ \mathbf{M}\ \frac{a\,(x\,(t))^2}{b} + o\,(1).$$

Подставляя это выражение в (14.280), находим плотность энтропии

$$h^{P/Q} = \frac{1}{2b} \mathbf{M} \left[a\,\big(x(t)\big)\right]^2, \qquad (14.281)$$

которая не зависит от *t* в стационарном случае.

Выше предполагалось, что *a* и *b* не зависят от времени *t* в соответствии с условием стационарности. Если бы это условие не было выполнено, мы получили бы описанным способом зависящую от времени энтропийную плотность

$$h^{P/Q}(t) = \frac{1}{2} \mathbf{M} \frac{a\,(x\,(t),\ t)^2}{b\,\big(x\,(t),\ t\big)} \qquad (14.282)$$

(условие независимости *b* от *x* здесь также может быть не выполнено). Через нее получаемая из (14.279) полная энтропия

$$H^{P/Q}_{x_0^T} = \mathbf{M} \ln \frac{P\left(dx_0^T\right)}{Q\left(dx_0^T\right)}$$

выражается интегрированием

$$H^{P/Q}_{x_0^T} = H^{P/Q}_{x(0)} + \int_0^T H_0^{P/Q}(t)\,dt. \qquad (14.283)$$

В стационарном же случае при b = const плотность энтропии может быть вычислена по формуле (14.281) при помощи стационарной плотности распределения $p_{ст}(x)$, удовлетворяющей стационарному уравнению Фоккера — Планка

$$-\frac{d}{dx}\left[a(x)\,p_{ст}(x)\right] + \frac{b}{2}\,\frac{d^2}{dx^2}\,p_{ст}(x) = 0.$$

Если ввести потенциальную функцию

$$U(x) = -\int_c^x a(x)\,dx, \qquad (14.284)$$

то стационарное распределение можно записать

$$p_{ст}(x) = \frac{1}{N}\,e^{-2U(x)/b}, \qquad N = \int e^{-2U(x)/b}\,dx, \quad (14.285)$$

и из (14.281) будем иметь

$$h^{P/Q} = \frac{1}{2bN}\int e^{-2U(x)/b}\left[\frac{dU(x)}{dx}\right]^2 dx.$$

Интегрированием по частям (при учете исчезновения плотности $p_{ст}(x)$ на границах, например, при $t \to \pm\infty$)эту формулу можно привести к виду

$$h^{P/Q} = \frac{1}{4N}\int e^{-2U(x)/b}\,\frac{d^2U(x)}{dx^2}\,dx = -\frac{1}{4}\,\mathbf{M}\,\frac{da(x)}{dx}. \quad (14.286)$$

Пример. Пусть функция $a(x)$ линейна: $a(x) = -\beta x + \gamma$. Чтобы процесс был стационарным, необходима положительность $\beta > 0$. Функция (14.284) имеет вид

$$U(x) = \frac{1}{2}\beta x^2 - \gamma x,$$

и распределение (14.285) является гауссовым:

$$p_{ст}(x) = \sqrt{\frac{\beta}{\pi b}}\,\exp\left[-\frac{\beta}{b}\left(x - \frac{\gamma}{\beta}\right)^2\right], \qquad N = \sqrt{\frac{\pi b}{\beta}}.$$

Применение формулы (14.286) дает

$$h^{P/Q} = \beta/4. \qquad (14.287)$$

Данный процесс является, как известно, гауссовым процессом, имеющим спектральную плотность $S(\omega) = 2b/(\omega^2 + \beta^2)$. Поэтому он совпадает с процессом, рассмотренным в п. 14.12 (пример 2). Результат (14.287), естественно, совпадает с соответствующим результатом (14.231), полученным на основе теории гауссовых процессов.

Приведенные в этом пункте результаты допускают обобщение на многомерный случай, когда имеется многокомпонентный марковский процесс $\{x(t)\} = \{x_1(t), \ldots, x_l(t)\}$. Пусть он характеризуется сносами $a_\rho(x, t)$, $\rho = 1, \ldots, l$, и матрицей локальных дисперсий $b_{\rho\sigma}(x, t)$, $\rho, \sigma = 1, \ldots, l$. Тогда выбирая меру *Q*, имеем производную Радона — Никодима

$$\frac{P(dx_0^T \mid x(0))}{Q(dx_0^T \mid x(0))} = \exp\left\{\int_0^T \sum_{\rho', \sigma'} a_{\rho'} b^{-1}_{\rho'\sigma'}\left(d^* x_{\sigma'} - \frac{1}{2} a_{\sigma'}\, dt\right)\right\}, \qquad (14.288)$$

что служит обобщением формулы (14.279). Здесь $b^{-1}_{\rho'\sigma'}$ —матрица, обратная невырожденной подматрице $b_{\rho'\sigma'}$, $\rho', \sigma' = 1, \ldots, l' \leqslant l$, матрице локальных дисперсий $b_{\rho\sigma}$, $\rho, \sigma = 1, \ldots, l$.

Тем же приемом, что и в одномерном случае, получаем из (14.288) плотность энтропии

$$h^{P/Q}(t) = \frac{1}{2}\, \mathrm{M} \sum_{\rho', \sigma'} a_{\rho'}\left(x(t), t\right) b^{-1}_{\rho'\sigma'}\left(x(t), t\right) a_{\sigma'}\left(x(t), t\right). \qquad (14.289)$$

В частности, если $l \times l$-матрица $b_{\rho\sigma}$ невырожденна, то суммирование в (14.289) по ρ' и σ' нужно проводить от 1 до *l*, а мера *Q* соответствует той же матрице локальных дисперсий $b_{\rho\sigma}$, но нулевому вектору сносов.

В стационарном случае усреднение в (14.289) означает интегрирование со стационарной плотностью распределения $p_{ст}(x)$, удовлетворяющей стационарному уравнению Фоккера — Планка

$$-\sum_\rho \frac{\partial}{\partial x_\rho}(a_\rho p_{ст}) + \frac{1}{2} \sum_{\rho, \sigma} \frac{\partial^2}{\partial x_\rho \partial x_\sigma}(b_{\rho\sigma} p_{ст}) = 0.$$

В важном частном случае стационарное распределение таково, что потоки вероятности обращаются в нуль:

$$G_\rho = a_\rho p_{ст} - \frac{1}{2} \sum_\sigma \frac{\partial}{\partial x_\sigma}(b_{\rho\sigma} p_{ст}) = 0.$$

В этом случае из (14.289) имеем

$$h^{P/Q}(t) = \frac{1}{4} \sum_{\rho', \sigma', \pi} \int a_{\rho'} b_{\rho'\sigma'}^{-1} \frac{\partial}{\partial x_\pi} [b_{\sigma'\pi} p_{\text{ст}}] \, dx_1 \ldots dx_l.$$

Применяя формулу Грина и учитывая исчезновение интеграла по границе, получаем

$$h^{P/Q}(t) = -\frac{1}{4} \sum_{\rho', \sigma', \pi} \int b_{\sigma'\pi} p_{\text{ст}} \frac{\partial}{\partial x_\pi} [a_{\rho'} b_{\rho'\sigma'}^{-1}] \, dx_1 \ldots dx_l =$$

$$= -\frac{1}{4} \mathbf{M} \, b_{\sigma'\pi} \frac{\partial}{\partial x_\pi} [a_{\rho'} b_{\rho'\sigma'}^{-1}].$$

Если матрица $b_{\rho\sigma}$ невырожденна и не зависит от *x,* то имеет место следующая формула:

$$h^{P/Q} = -\frac{1}{4} \sum_{\rho, \sigma, \pi} \mathbf{M} \, b_{\sigma\pi} b_{\sigma\rho}^{-1} \frac{\partial a_\rho}{\partial x_\pi} =$$

$$= -\frac{1}{4} \mathbf{M} \sum_{\rho} \frac{\partial a_\rho}{\partial x_\rho} \equiv -\frac{1}{4} \mathbf{M} \operatorname{div} a,$$

которая является многомерным обобщением формулы (14.286).

14.15. Энтропия комбинированного марковского процесса, условного процесса и части компонент марковского процесса

1. На случай непрерывного времени могут быть обобщены результаты и методы, изложенные ранее. Предполагается, что совокупный процесс $\{\xi(t)\} = \{x(t), \; y(t)\}$ является марковским. Теория марковских процессов позволяет вычислять энтропию

$$H^{P/Q}_{\xi_a^T} = H^{P/Q}_{x_a^T y_a^T},$$

где Q — вероятностная мера, относительно которой мера P является абсолютно непрерывной. Меру Q удобно подбирать таким образом, чтобы процессы $\{x(t)\}$ и $\{y(t)\}$ были для этой меры независимыми:

$$Q(d\xi_a^T) = Q_1(dx_a^T) \, Q_2(dy_a^T)$$

и марковскими, т. е. чтобы

$$Q(dx_t^{t+\tau} \mid x_a^t y_a^t) = Q(dx_t^{t+\tau} \mid x(t)), \quad Q(dy_t^{t+\tau} \mid x_a^t y_a^t) =$$

$$= Q(dy_t^{t+\tau} \mid y(t)).$$

Тогда будут справедливы соотношения

$$H^{P/Q}_{x_a^T y_a^T} = H^{P/Q_2}_{y_a^T} + H^{P/Q_1}_{x_a^T \mid y_a^T}, \qquad (14.290)$$

$$H^{P/Q}_{x_a^T y_a^T} = H^{P/Q_1}_{x_a^T} + H^{P/Q_2}_{y_a^T | x_a^T}, \tag{14.291}$$

$$H^{P/Q}_{x_t^{t+\tau} y_t^{t+\tau} | x_a^t y_a^t} = H^{P/Q}_{x_t^{t+\tau} y_t^{t+\tau} | x(t)\, y(t)}, \tag{14.292}$$

аналогичные соотношениям (14.104), (14.109), (14.85) дискретной версии.

В случае непрерывного времени удобно ввести энтропийные плотности

$$h^{P/Q}_{xy}(t) = \lim_{\tau\to 0} \frac{1}{\tau} H^{P/Q}_{x_t^{t+\tau} y_t^{t+\tau} | x(t)\, y(t)}, \tag{14.293}$$

$$h^{P/Q_2}_{y}(t) = \lim_{\tau\to 0} \frac{1}{\tau} H^{P/Q_2}_{y_t^{t+\tau} | y_a^t}, \tag{14.294}$$

$$h^{P/Q_1}_{x|y}(t) = \lim_{\tau\to 0} \frac{1}{\tau} \left[H^{P/Q_1}_{x_a^{t+\tau} | y_a^{t+\tau}} - H^{P/Q_1}_{x_a^t | y_a^t} \right]. \tag{14.295}$$

Полагая в (14.290) сначала $T = t + \tau$, а затем $T = t$, взяв разность этих выражений и используя свойство аддитивности, получаем соотношение

$$H^{P/Q}_{x_t^{t+\tau} y_t^{t+\tau} | x_a^t y_a^t} = H^{P/Q_2}_{y_t^{t+\tau} | y_a^t} + H^{P/Q_1}_{x_a^{t+\tau} | y_a^{t+\tau}} - H^{P/Q_1}_{x_a^t | y_a^t}. \tag{14.296}$$

Поделив (14.296) на τ и перейдя к пределу $\tau \to 0$, получим соотношение

$$h^{P/Q}_{xy}(t) = h^{P/Q_2}_{y}(t) + h^{P/Q_1}_{x|y}(t). \tag{14.297}$$

Аналогичное соотношение

$$h^{P/Q}_{xy}(t) = h^{P/Q_1}_{x}(t) + h^{P/Q_2}_{y|x}(t) \tag{14.298}$$

справедливо, очевидно, и для другой пары энтропийных плотностей

$$\begin{gathered} h^{P/Q}_{x}(t) = \lim_{\tau\to 0} \frac{1}{\tau} H^{P/Q_1}_{x_t^{t+\tau} | x_a^t}, \\ h^{P/Q_2}_{y|x}(t) = \lim_{\tau\to 0} \frac{1}{\tau} \left[H^{P/Q_2}_{y_a^{t+\tau} | x_a^{t+\tau}} - H^{P/Q_2}_{y_a^t | x_a^t} \right]. \end{gathered} \tag{14.293}$$

В стационарном случае плотность $h^{P/Q}_{xy}$, как видно из (14.293), не зависит от *t*. Другие энтропийные плотности при этом целесообразно определить с помощью дополнительного предельного перехода $a \to -\infty$. В этом случае энтропии будут строго пропорциональны τ и предельный переход $\tau \to 0$ станет излишним. Формулы (14.294), (14.295) заменяются следующими:

$$h_y^{P/Q_2}(t) = \frac{1}{\tau} H^{P/Q_2}_{y_t^{t+\tau} | y_{-\infty}^t},$$

$$h_{x|y}^{P/Q_1}(t) = \frac{1}{\tau} \lim_{a \to -\infty} \left[H^{P/Q_1}_{x_a^{t+\tau} | y_a^{t+\tau}} - H^{P/Q_1}_{x_a^t | y_a^t} \right]$$

(и аналогично для другой пары h_x^{P/Q_1}, $h_{y/x}^{P/Q_2}$). При этом соотношения (14.297), (14.299) сохраняют свое значение. Все эти энтропийные плотности в стационарном случае будут постоянными. Можно доказать, что они совпадают с удельными энтропиями, т. е. доказать равенства:

$$h_y^{P/Q_2} = \lim_{t \to \infty} \frac{1}{t} H^{P/Q_2}_{y_0^t}, \quad h_{x|y}^{P/Q_1} = \lim_{t \to \infty} \frac{1}{t} H^{P/Q_1}_{x_0^t | y_0^t}.$$

Можно доказать кроме того, что для плотности $h_{x|y}^{P/Q_1}$ (и аналогично для $h_{y|x}^{P/Q_2}$) в стационарном случае справедлива формула

$$h_{x|y}^{P/Q_1} = \frac{1}{\tau} H^{P/Q_1}_{x_0^\tau | x(0)\, y_0^\infty} = \frac{1}{\tau} H^{P/Q_1}_{x_0^\tau | x(\tau)\, y_{-\infty}^\tau} \quad (\tau > 0 \text{ любое}).$$

Все эти утверждения распространяют на случай непрерывного времени.

Обобщая методы изложенные ранее, можно вычислить энтропию $H^{P/Q_2}_{y_0^T}$ или ее плотность h_y^{P/Q_2} для части компонент *{y}* марковского процесса ξ, и энтропию $H^{P/Q_1}_{x_0^T | y_0^T}$ или $h_{x|y}^{P/Q_1}$ условного марковского процесса $\{x(t)\}$, а также аналогичные энтропии $H^{P/Q_1}_{x_0^T}$, h_x^{P/Q_1} и

$$H^{P/Q_2}_{y_0^T | x_0^T}, \quad h_{y|x}^{P/Q_2}.$$

Рассмотрим энтропию

$$H^{P/Q_2}_{y_t^{t+\tau} | y_0^t}.$$

Ее можно представить в таком виде:

$$H^{P/Q_2}_{y_t^{t+\tau} | y_0^t} = \mathbf{M} \ln \frac{P\left[dy_t^{t+\tau} \mid y_0^t\right]}{Q_2\left[dy_t^{t+\tau} \mid y(t)\right]}. \tag{14.300}$$

Тогда имеем

$$\frac{P\left[dy_t^{t+\tau} \mid y_0^t\right]}{Q_2\left[dy_t^{t+\tau} \mid y(t)\right]} = \mathbf{M}\left\{\frac{P\left[dy_t^{t+\tau} \mid x(t),\ y_0^t\right]}{Q_2\left[dy_t^{t+\tau} \mid y(t)\right]} \middle| y_0^t\right\}.$$

В силу условия Маркова

$$P\left[dy_t^{t+\tau} \mid x(t),\ y_0^t\right] = P\left[dy_t^{t+\tau} \mid x(t),\ y(t)\right]$$

указанное равенство можно записать так:

$$\frac{P\left[dy_t^{t+\tau} \mid y_0^t\right]}{Q_2\left[dy_t^{t+\tau} \mid y(t)\right]} = \int \frac{P\left[dy_t^{t+\tau} \mid x(t),\ y(t)\right]}{Q_2\left[dy_t^{t+\tau} \mid y(t)\right]} W\left(dx(t)\right), \quad (14.301)$$

где обозначено

$$W\left[dx(t)\right] = P\left[dx(t) \mid y_0^t\right].$$

Подставляя (14.301) в (14.300), получаем формулу

$$H^{P/Q_2}_{y_t^{t+\tau} \mid y_0^t} = \mathbf{M} \ln\left[\int \frac{P\left[dy_t^{t+\tau} \mid x(t),\ y(t)\right]}{Q_2\left[dy_t^{t+\tau} \mid y(t)\right]} W(dx)\right]. \quad (14.302)$$

Если апостериорную меру W трактовать как распределение в комбинированном пространстве $\xi = (x,\ y)$ (см. (14.120)), то формула (14.202) запишется в виде

$$H^{P/Q_2}_{y_t^{t+\tau} \mid y_0^t} = \mathbf{M} \ln\left[\int \frac{P\left(dy_t^{t+\tau} \mid \xi\right)}{Q_2\left(dy_t^{t+\tau} \mid y\right)} W(d\xi)\right],$$

или

$$H^{P/Q_2}_{y_t^{t+\tau} \mid y_0^t} = \mathbf{M} \int P\left(dy_t^{t+\tau} \mid \xi\right) W(d\xi) \ln\left[\frac{P\left(dy_t^{t+\tau} \mid \xi\right)}{Q_2\left(dy_t^{t+\tau} \mid y\right)} W(d\xi)\right], \quad (14.303)$$

где усреднение **M** проводится только по случайным переменным $W(\bullet)$, образующим вторичный апостериорный W-процесс, являющийся марковским процессом с известными вероятностями перехода.

Формулы (14.302), (14.303) являются обобщением формул (14.118), (14.119) дискретной версии. Они справедливы при любых τ, но для вычисления h_y^{P/Q_2} ими удобнее пользоваться при малых τ.

Это будет видно из рассматриваемых ниже частных случаев.

2. Пусть теперь $\{x(t)\}$ — марковский процесс с дискретным числом состояний, подобный процессу, рассмотренному в п. 14.13. Он характеризуется введенной в п. 14.13 дифференциальной матрицей перехода $\pi_t(x, x')$, определяющей вероятности перехода $P(\xi(t + \Delta) \mid \xi(t))$. Выбирая в качестве Q_1 пуассоновскую меру для точек перехода, получаем плотносп $h_x^{P/Q_1}(t)$ энтропии $H^{P/Q_1}_{x_0^T}$

$$h_x^{P/Q_1}(t)=\sum_{x_t}P\left(x(t)\right)\Big\{1+$$
$$+\sum_{x'\neq x_t}\left[\ln\pi_t\left(x(t),\,x'\right)-1\right]\pi_t\left(x(t),\,x'\right)\Big\}. \quad (14.304)$$

Процесс $\{y(t)\}=\{y_1(t),\,\ldots,\,y_l(t)\}$ пусть строится следующим образом. Задан зависящий не только от *y (t),* но и от *x (t)* вектор сносов $a_\rho(x(t),\; y(t),\; t)$, а также матрица локальных дисперсий $b_{\rho\sigma}(y(t),\; t)$. Последнюю предполагаем невырожденной и не зависящей от *x (t).*

Тогда при фиксированной реализации *{x(t)}* процесс *y(t)* будет диффузионным процессом, рассмотренным в п. 14.14. Применяя полученные там результаты, мы можем найти плотность $h_{y|x}^{P/Q_2}(t)$ энтропии $H^{P/Q_2}_{y_0^T|x_0^T}$. В качестве меры Q_2 выбираем меру диффузионного процесса с нулевыми сносами и с той же матрицей локальных дисперсий $b_{\rho\sigma}(y,\; t)$. В соответствии с (14.302) имеем

$$h_y^{P/Q_2}(t\,|\,x_0^T)=\frac{1}{2}\sum_{\rho,\,\sigma}\int a_\rho\left(x(t),\,y(t),\,t\right)b_{\rho\sigma}^{-1}\left(y(t),\,t\right)\times$$
$$\times\, a_\sigma\left(x(t),\;y(t),\;t\right)P\left(dy(t)\,|\,x_0^T\right). \quad (14.305)$$

Для получения плотности $h_{y|x}^{P/Q_2}$ энтропии $H^{Q/P_2}_{y_0^T|x_0^T}$ остается в (14.305) произвести дополнительное усреднение по x_0^T:

$$h_{y|x}^{P/Q_2}(t)=\frac{1}{2}\sum_{\rho,\,\sigma}a_\rho\left(x(t),\,y(t),\,t\right)b_{\rho\sigma}^{-1}\left(y(t),\,t\right)\times$$
$$\times\, a_\sigma\left(x(t),\;y(t),\;t\right)P\left(dy(t)\;dx(t)\right). \quad (14.306)$$

В силу аддитивности (14.291), (5.11.9) тем самым найдена энтропийная плотность $h_{xy}^{P/Q}$ для комбинированного марковского процесса $\{\xi(t)\}=\{x(t),\;y(t)\}$. Именно, в (14.299) следует подставить выражения (14.304), (14.306).

Другие энтропийные плотности h_y^{P/Q_2} и $h_{x|y}^{P/Q_1}$ согласно (14.290) связаны аналогичным соотношением (14.297). Поэтому для завершения вычисления всех плотностей остается вычислить одну из них.

Описанный выше процесс $\{y(t)\}$ (при нефиксированной реализации *{x(t)})* является немарковским процессом. Соответствующую ему энтропию вычислим, используя формулу (14.302).

При малых τ из (14.288) имеем

$$P\left[dy_t^{t+\tau}\,|\,x(t),\,y(t)\right]/Q_2\left[dy_t^{t+\tau}\,|\,y(t)\right]=\exp\left\{\sum_{\rho,\sigma} a_\rho b_{\rho\sigma}^{-1}\times\right.$$
$$\left.\times\left[y_\sigma(t+\tau)-y_\sigma(t)-\frac{\tau}{2}a_\sigma\right]+o(\tau)\right\}=1+\sum_{\rho,\,\sigma} a_\rho b_{\rho\sigma}^{-1}\times$$
$$\times\left[y_\sigma(t+\tau)-y_\sigma(t)-\frac{\tau}{2}a_\sigma\right]+\frac{1}{2}\sum a_\rho b_{\rho\sigma}^{-1}\left[y_\sigma(t+\tau)-\right.$$
$$\left.-y_\sigma(t)\right]\left[y_\pi(t+\tau)-y_\pi(t)\right]b_{\pi\nu}^{-1}a_\nu+o(\tau).$$

Подставляя это выражение в (14.301) и для краткости обозначая $\sum\limits_{x(t)}\ldots\ W_t(x_t)=\mathbf{M}_{ps}\ \ldots$, будем иметь

$$P\left[dy_t^{t+\tau}\,|\,y_0^t\right]/Q_2\left[y_t^{t+\tau}\,|\,y(t)\right]=1+\sum_{\rho,\,\sigma}(\mathbf{M}_{ps}\,a_\rho)\,b_{\rho\sigma}^{-1}\times$$
$$\times\left[y_\sigma(t+\tau)-y_\sigma(t)\right]+\frac{1}{2}\sum\left\{b_{\rho\sigma}^{-1}\left[y_\sigma(t+\tau)-y_\sigma(t)\right]\times\right.$$
$$\left.\times\left[y_\pi(t+\tau)-y_\pi(t)\right]b_{\pi\nu}^{-1}-\tau b_{\rho\nu}^{-1}\right\}\mathbf{M}_{ps}(a_\nu a_\rho)+o(\tau).$$

Согласно (14.300), (14.302) следует взять логарифм от последнего выражения. При его логарифмировании нужно оставить лишь следующие члены:

$$\ln\frac{P\left[dy_t^{t+\tau}\,|\,y_0^t\right]}{Q_2\left[dy_t^{t+\tau}\,|\,y(0)\right]}=(\mathbf{M}_{ps}\,a_\rho)\,b_{\rho\sigma}^{-1}\left[y_\sigma(t+\tau)-y_\sigma(t)\right]+$$
$$+\frac{1}{2}\left\{b_{\rho\sigma}^{-1}\left[y_\sigma(t+\tau)-y_\sigma(t)\right]\left[y_\pi(t+\tau)-y_\pi(\tau)\right]b_{\pi\nu}^{-1}-b_{\rho\nu}^{-1}\right\}\times$$
$$\times\mathbf{M}_{ps}(a_\nu a_\rho)-\frac{1}{2}\left\{(\mathbf{M}_{ps}\,a_\rho)\,b_{\rho\sigma}^{-1}\left[y_\sigma(t+\tau)-y_\sigma(t)\right]\right\}^2+o(\tau).\qquad(14.307)$$

Усреднение его будем проводить в несколько этапов: сначала усредним по $y_t^{t+\tau}$ при фиксированных $x(t),\ y_0^t$, потом по *x(t)* (с весом $W_t\ (x(t))$) при фиксированных y^t_0 и в последнюю очередь по y^t_0 или, что то же самое, по *W*. На первом этапе усреднения нужно учесть формулы

$$\mathbf{M}\left\{y_\sigma(t+\tau)-y_\sigma(t)\,|\,x(t),\ y(t)\right\}=a_\sigma\left(x(t),\,y(t),\,t\right)\tau+o(\tau),$$
$$\mathbf{M}\left\{\left[y_\sigma(t+\tau)-y_\sigma(t)\right]\left[y_\pi(t+\tau)-y_\pi(t)\right]\,|\,x(t),\,y(t)\right\}=$$
$$=b_{\sigma\pi}\left(x(t),\,y(t),\,t\right)\tau+o(\tau),$$

так что

$$\frac{1}{\tau}\mathbf{M}\left[\ln\frac{P\left[dy_t^{t+\tau}\mid y_0^t\right]}{Q_2\left[dy_t^{t+\tau}\mid y(0)\right]}\middle|\, x(0),\, y_0^t\right]=(\mathbf{M}_{\text{ps}}\,a_\rho)\,b_{\rho\sigma}^{-1}\,a_\sigma-$$

$$-\frac{1}{2}(\mathbf{M}_{\text{ps}}\,a_\rho)\,b_{\rho\sigma}^{-1}(\mathbf{M}_{\text{ps}}\,a_\sigma)+o(1).$$

Дальнейшее усреднение приведет к результату

$$\frac{1}{\tau}H^{P/Q_2}_{y_t^{t+\tau}\mid y_0^t}=\frac{1}{2}\mathbf{M}(\mathbf{M}_{\text{ps}}\,a_\rho)\,b_{\rho\sigma}^{-1}(\mathbf{M}_{\text{ps}}\,a_\sigma)+o(1).$$

Совершая предельный переход $\tau\to 0$, найдем

$$h_y^{P/Q_2}(t)=\frac{1}{2}\mathbf{M}\sum_{\rho,\,\sigma}(\mathbf{M}_{\text{ps}}\,a_\rho)\,b_{\rho\sigma}^{-1}(\mathbf{M}_{\text{ps}}a_\sigma)=$$

$$=\frac{1}{2}\int P\left(dW_t,\, dy(t)\right)\sum_{x,\,x'}\sum_{\rho,\,\sigma}W_t(x)\,a_\rho\left(x,\, y(t),\, t\right)\times$$

$$\times\, b_{\rho\sigma}^{-1}\left(y(t),\, t\right)a_\sigma\left(x',\, y(t),\, t\right)W_t(x'). \tag{14.308}$$

Найти входящее сюда распределение $P(dW_t,\ dy(t))$ помогает то обстоятельство, что переменные $\{W_t,\ y_t\}$ образуют марковский процесс, названный Стратоновичем вторичным апостериорным *W*-процессом. Там найдено уравнение Фоккера — Планка, которому удовлетворяет плотность распределения вероятостей $p_t(W,\ y)$ процесса $\{W_t,\ y\ (t)\}$. Оно имеет вид

$$\frac{dp_t}{dt}=\frac{\partial}{\partial W(x')}\left[\pi_{xx'}\,W(x)\,p_t\right]-\frac{\partial}{\partial y_\rho}\left[(\mathbf{M}_{\text{ps}}\,a_\rho)\,p_t\right]+$$

$$+\frac{1}{2}\sum\frac{\partial^2}{\partial W(x)\,\partial W(x')}\left\{W(x)\left[a_\rho(x)-\mathbf{M}_{\text{ps}}\,a_\rho\right]b_{\rho\sigma}^{-1}\times\right.$$

$$\left.\times\left[a_\sigma(x')-\mathbf{M}_{\text{ps}}\,a_\sigma\right]W(x')\,p_t\right\}+\sum\frac{\partial^2}{\partial W(x)\,\partial y_\rho}\left\{W(x)\times\right.$$

$$\left.\times\left[a_\rho(x)-\mathbf{M}_{\text{ps}}\,a_\rho\right]p_t\right\}+\frac{1}{2}\sum\frac{\partial^2}{\partial y_\rho\partial y_\sigma}\left[a_{\rho\sigma}\,p_t\right]. \tag{14.309}$$

Указанные результаты справедливы как в стационарном, так и в нестационарном случае. В стационарном случае функции $\pi_t(x,\ x'),\ \ a_\rho\ (x,\ y,\ t),\ b_{\rho\sigma}\ (y,\ t)$ не зависят от времени. Плотность энтропии h_y^{P/Q_2} при этом постоянна и вычисляется по формуле (14.308) при помощи стационарного распределения $P_{\text{ст}}\ (dW_t, dy\ (t))$, являющегося решением стационарного уравнения (14.309), в котором следует положить $dp_t/dt = 0$.

По аналогии с (14.308) вместо *W (x)* можно рассматривать *W (•)* как распределение $W_t(dxdy)$ во всем «фазовом пространстве» $\xi = (x, y)$ марковского процесса (такой процесс $\{W_t(.)\}$ будет марковским). Тогда формулу (14.308) можно будет записать следующим образом:

$$h_y^{P/Q_2}(t) = \frac{1}{2}\int P(dW_t)\int\int\sum_{\rho,\sigma} W_t(d\xi)\,a_\rho(\xi, t)\times$$

$$\times\, b_{\rho\sigma}^{-1}(y, t)\,a_\sigma(\xi', t)\,W_t(d\xi').$$

Итак, удельная энтропия h_y^{P/Q_2}, а значит и энтропия $h_{x|y}^{P/Q_2} = h_{xy}^{P/Q_2} - h_y^{P/Q_2}$ вычисляется методами, использующими результаты теории условных марковских процессов.

Пример. Пусть *{x(t)}* — процесс с двумя состояниями и матрицей перехода (14.275). В п. 14.13 была найдена соответствующая ему удельная энтропия

$$h_x^{P/Q_1} = 1 + \frac{\mu\nu}{\mu+\nu}\ln\frac{\mu\nu}{e^2}. \tag{14.310}$$

Процесс *{y(t)}* определим как одномерный диффузионный процесс с постоянной локальной дисперсией *b* и с не зависящими от *у* и *t*, но зависящими от *x* сносами *а (х).*

Тогда формула (14.306) дает

$$h_{y|x}^{P/Q_2} = \frac{1}{2b}\left[\mathbf{P}(x=1)\,a^2(1) + \mathbf{P}(x=2)\,a^2(2)\right],$$

или

$$h_{y|x}^{P/Q_2} = \frac{\nu a^2(1) + \mu a^2(2)}{2b(\mu+\nu)}, \tag{14.311}$$

если учесть (14.276).

Сумма выражений (14.310), (14.311) дает удельную энтропию $h_{xy}^{P/Q}$ комбинированного процесса.

Перейдем к вычислению удельной энтропии h_u^{P/Q_2} для данного примера. Вводя переменную $z_t = W_t(1) - W_t(2)$, средний снос $\mathbf{M}_{ps}\,a$ можно представить в виде

$$\mathbf{M}_{ps}\,a = W_t(1)\,a(1) + W_t(2)\,a(2) = \frac{a(1)+a(2)}{2} + \frac{a(1)-a(2)}{2}\,z.$$

Подставляя это выражение в (14.308), имеем

$$h_y^{P/Q_2} = \frac{1}{8b}\left[a(1)+a(2)\right]^2 + \frac{1}{4b}\mathbf{M}\left\{\left[a^2(1)-a^2(2)\right]z + \right.$$
$$\left. + \frac{1}{2}\left[a(1)-a(2)\right]^2 z^2\right\}. \qquad (14.312)$$

Нетрудно понять, что усреднение апостериорных вероятностей $W_t(x) \;\; = \mathbf{P}\left[x_t = x \mid y_0^t\right]$ приводит к априорным вероятностям $\mathbf{P}\left[x_t = x\right]$, которые в стационарном случае имеют вид (14.276). Поэтому

$$\mathbf{M}z = \mathbf{M}\left[W_t(1)-W_t(2)\right] = \mathbf{P}\left[x=1\right] - \mathbf{P}\left[x=2\right] = \frac{\nu-\mu}{\mu+\nu}$$

и формулу (14.312) можно преобразовать к виду

$$h_y^{P/Q_0} = \frac{1}{8b}\left[a(1)+a(2)\right]^2 + \frac{(\nu-\mu)\left[a^2(1)-a^2(2)\right]}{4b(\mu+\nu)} +$$
$$+ \frac{1}{8b}\left[a(1)-a(2)\right]^2 \int_{-1}^{1} z^2 p_{\text{ст}}(z)\,dz$$

или

$$h_y^{P/Q_2} = \frac{\nu a^2(1)+\mu a^2(2)}{2b(\mu+\nu)} - \frac{\left[a(1)-a(2)\right]^2}{8b}\int_{-1}^{1}(1-z^2)\,p_{\text{ст}}(z)\,dz. \qquad (14.313)$$

Процесс $\{z_t\}$ в данном случае является марковским сам по себе.

Ему соответствует уравнение Фоккера — Планка

$$\dot{p}(z) = -\frac{\partial}{\partial z}\left\{\left[\nu-\mu+(\mu+\nu)z\right]p(z)\right\} +$$
$$+ \frac{1}{2b}\frac{\partial^2}{\partial z^2}\left[(1-z^2)^2 p(z)\right],$$

вытекающее из (14.309).

Приравнивая производную $\dot{p}_{\text{ст}}(z)$ нулю и интегрируя получающееся уравнение, получаем стационарную плотность распределения вероятностей

$$p_{\text{ст}}(z) = \frac{\text{const}}{(1-z^2)^2}\exp\left\{2b\int_0^z\left[\nu-\mu-(\mu+\nu)x\right]\frac{dx}{(1-x^2)^2}\right\}. \qquad (14.314)$$

Входящий в выражение (14.313) интеграл с этой плотностью распределения был вычислен и оказался равным

$$1-\int z^2 p(z)\,dz = 4K_q\left(b\sqrt{\mu\nu}\right)\left\{2K_q\left(b\sqrt{\mu\nu}\right) + \right.$$
$$\left. + \sqrt{\frac{\nu}{\mu}}K_{q+1}\left(b\sqrt{\mu\nu}\right) + \sqrt{\frac{\mu}{\nu}}K_{q-1}\left(b\sqrt{\mu\nu}\right)\right\}^{-1}, \qquad (14.315)$$

где $q = \frac{b}{2}(\nu-\mu)$, а $K_q(z)$ — функция Макдональда.

Подстановка (14.315) в (14.313) решает задачу вычисления удельной энтропии h_y^{P/Q_2} стационарного немарковского процесса *у*.

Комбинируя (14.310),. (14.311), (14.313), легко найти энтропию

$$h_{x|y}^{P/Q_1} = h_x^{P/Q_1} + h_{y|x}^{P/Q_2} - h_y^{P/Q_2}$$

условного марковского процесса *х (t)*.

3. Изложенный выше вывод формулы (14.308) применим и в других случаях, например, когда процесс *{x(t)}* является диффузионным марковским процессом или образует часть компонент комбинированного диффузионного марковского процесса $\{\xi(t)\} = \{x(t),\ y(t)\}$. Рассмотрим последний случай подробнее. Пусть комбинированный марковский процесс, имеющий компоненты

$$\xi_\alpha(t) = x_\alpha(t), \quad \alpha = 1,\dots, m; \qquad \xi_\rho(t) = y_\rho(t),\ \rho = m+1,\dots, m+l,$$

описывается вектором сносов $a_j\ (\xi,\ t)$ и $(m+l) \times (m+l)$-матрицей локальных дисперсий $b_{jk}\ (\xi,\ t),\ j,\ k = 1,\ \dots,\ m+l$. Предполагается, что $l\times l$ — подматрица последней $b_{\rho\sigma}\ (\xi,\ t)$, ρ, $\sigma = m+1,\ \dots,\ m+l$ является невырожденной и не зависящей от *х*.

Тогда по аналогии с (14.308) справедлива формула

$$h_y^{P/Q_2}(t) = \frac{1}{2}\mathbf{M} \sum_{\rho,\ \sigma=m+1}^{m+l} (\mathbf{M}_{\text{ps}}\, a_\rho)\, b_{\rho\sigma}^{-1}\, (\mathbf{M}_{\text{ps}}\, a_\sigma), \quad (14.316)$$

где усреднение $\mathbf{M}_{\text{ps}}$ соответствует усреднению с апостериорной плотностью вероятности

$$\mathbf{M}_{\text{ps}}\, a_\sigma = \int a_\sigma(x,\ y,\ t)\, w_t(x)\, dx_1 \dots\ dx_m,$$

которая удовлетворяет уравнению

$$\frac{\partial}{\partial t}\, w_t(x) = \frac{1}{2} \sum_{\alpha,\ \beta} \frac{\partial^2}{\partial x_\alpha\, \partial x_\beta} [b_{\alpha\beta}\, w_t] -$$

$$- \sum_{\alpha=1}^{m} \frac{\partial}{\partial x_\alpha}\left[a_\alpha w_t + \sum_{\sigma,\rho=m+1}^{m+l} b_{\alpha\rho}\, b_{\rho\sigma}^{-1}\, w_t\left(\frac{d^* y_\sigma}{dt} - \mathbf{M}_{\text{ps}}\, a_\sigma\right)\right] +$$

$$+ \sum_{\rho,\ \sigma=m+1}^{m+l} (a_\rho - \mathbf{M}_{\text{ps}}\, a_\rho)\, b_{\rho\sigma}^{-1}\, w_t \left[\frac{d^* y_\sigma}{dt} - \mathbf{M}_{\text{ps}}\, a_\sigma\right].$$

Второе усреднение в (14.316) соответствует усреднению по марковскому процессу $\{w_t(x),\ \ y(t)\}$ как стационарному случайному процессу.

4. До сих пор процесс *{y(t)}* предполагался диффузионным Аналогичные результаты можно получить, скажем для того случая, когда процесс *{y (t)}* есть процесс с конечным или счетным числом состояний.

Пусть *{x{t)}* — марковский процесс с матрицей перехода $\pi_t\ (x,\ x')$, подобный процессу в п. 2. При фиксированной реализации *{x(t)}* процесс *{y(t)}* пусть является марковским процессом с дискретными состояниями, описываемыми дифференциальными вероятностями перехода $\pi_t\ (x,\ y,\ y')$, зависящими от $x = x\ (t)$.

Тогда комбинированный процесс $\{\xi(t)\} = \{x(t),\ y(t)\}$ также будет марковским процессом с дискретными состояниями. Ему будет соответствовать дифференциальная матрица перехода

$$\pi_t(\xi,\xi') = \pi_t(x,x')\,\delta_{yy'} + \pi_t(x,y,y')\,\delta_{xx'}. \qquad (14.317)$$

В данном случае энтропия процесса *{x(t)}* и энтропия комбинированного процесса *{x(t), y(t)}* могут быть подсчитаны путем применения формулы (14.273). Подсчет энтропии $H^{P/Q_2}_{y_0^T}$ или ее плотности $h_y^{P/Q_2}\ (t)$ требует особого рассмотрения, поскольку процесс *{y (t)},* взятый в отдельности, является немарковским.

В принципиальном отношении соответствующий расчет может быть проведен так же, как и в п. 2. Воспользуемся формулой (14.302), выбирая теперь в качестве меры Q_2 пуассоновскую меру (для моментов перескока процесса *y(t))* с единичной плотностью. При малом τ будем иметь

$$\frac{P\left[dy_t^{t+\tau}\,\middle|\,x(t),\,y(t)\right]}{Q_2\left[dy_t^{t+\tau}\,\middle|\,y(t)\right]} =$$

$$= \begin{cases} e^{\tau}\left[1+\pi_t\big(x(t),\,y(t),\,y(t)\big)\tau\right] + O(\tau^2) & \text{при } y(t+\tau)=y(t),\\ e^{\tau}\,\pi_t\big(x(t),\,y(t),\,y(t+\tau)\big)\tau + O(\tau^2) & \text{при } y(t+\tau)\neq y(t). \end{cases}$$

Подставляя это выражение в (14.302), находим

$$H^{P/Q_2}_{y_t^{t+\tau}\,|\,y_0^t} = \mathbf{M}\left[1+\mathbf{M}_{\text{ps}}\,\pi_t\big(x(t),\,y(t),\,y(t)\big)\tau\right]\ln\Big\{e^{\tau}\Big[1+$$

$$+\mathbf{M}_{\text{ps}}\,\pi_t\big(x(t),\,y(t),\,y(t)\big)\tau\Big]\Big\} + \sum_{y(t+\tau)\neq y(t)}\mathbf{M}_{\text{ps}}\,\pi_t\big(x(t),\ y(t),$$

$$y(t+\tau)\big)\,\tau\ln\left[e^{\tau}\mathbf{M}_{\text{ps}}\,\pi_t\big(x(t),\,y(t),\,y(t+\tau)\big)\right] + O(\tau^2). \qquad (14.318)$$

Здесь, в первую очередь, произведено усреднение $\mathbf{M}\ [\ldots\,|\,x(t),$ $y_0^t\ (t)]$ по $y_t^{t+\tau}$, затем усреднение $\mathbf{M}\ [\ldots\,|\,y_0^t]$ по *x (t)* с весом W_t, а потом усреднение по прочим переменным.

Из (14.318) вытекает следующая формула для энтропийной плотности:

$$h_y^{P/Q_2}(t) = \mathbf{M}\Big\{1 + \sum_{y' \neq y(t)} \Big[\mathbf{M}_{ps}\,\pi\big(x(t), y(t), y'\big)\Big] \times$$

$$\times \Big[\ln\Big[\mathbf{M}_{ps}\,\pi\big(x(t), y(t), y'\big)\Big] - 1\Big]\Big\}. \qquad (14.319)$$

Отличие этой формулы от формулы

$$h_y^{P/Q_2} = \mathbf{M}\Big\{1 + \sum_{y' \neq y(t)} \pi\big(y(t), y'\big)\Big[\ln \pi\big(y(t), y'\big) - 1\Big]\Big\}, \qquad (14.320)$$

(т. е. от формулы (14.273), взятой при $\xi = y$) в том, что $\pi\ (y,\ y')$ заменены на апостериорные средние

$$\mathbf{M}_{ps}\,\pi\big(x(t), y(t), y'\big) = \sum_x W_t(x)\,\pi\big(x, y(t), y'\big)$$

и имеется усреднение не по $y\ (t)$ с весом $P(dy)$, а усреднение по $\{W_t,\ y\ (t)\}$ с весом $P\ [dW_t dy\ \ (t)]$. Процесс $\{W_t,\ y\ (t)\}$ является вторичным апостериорным марковским процессом и для него нетрудно найти вероятности перехода, которые определяют, в частности, стационарное распределение $P_{ст}\ [dW_t\ dy(t)\}$.

Формулу (14.319) интересно сравнить с условной энтропией

$$h_{y|x}^{P/Q_2}(t) = \mathbf{M}\Big\{1 + \sum_{y' \neq y(t)} \pi\big(x(t), y(t), y'\big) \times$$

$$\times\Big[\ln \pi\big(x(t), y(t), y'\big) - 1\Big]\Big\}. \qquad (14.321)$$

Это выражение мы записали по аналогии с (14.320), считая значение *x (t)* фиксированным, а затем усредняя по нему.

Изложенный метод вычисления плотности энтропии h_y может быть распространен и на случай, когда процесс *{x (t)}* в отдельности не является марковским, а марковским процессом (с дискретными состояниями) является комбинированный процесс *{x(t), y(t)}*. Пусть он описывается дифференциальной матрицей перехода $\pi_t\ (x,\ y,\ x',\ y')$. В этом случае вид результирующей формулы (14.319) остается без изменения. Аналогично вычисляется и $h_x^{P/Q}$.

Из вышеизложенного видно, что описанный метод вычисления энтропии для части компонент марковского процесса имеет широкую область применения. В нем наиболее трудным этапом является отыскание распределения $P\,(dW_t,\ dy\,(t))$ для вторичного апостериорного *W*-процесса.

14.15. Оценка нечеткости. Нечеткая энтропия

Большинство вещей в нашем мире в той или иной степени являются неопределенными, неточными, неполными, качественными. Все эти характеристики очень трудно оценивать используя классические математические методы. Измерение этой нечеткости является предусловием анализа сложных систем. Для этого Шэннон ввел понятие "энтропия", которое позволяет описать степень нечеткости в случайных данных. Когда мы говорим об энтропии в терминах нечеткой логики, мы получаем нечеткую энтропию, которая описывает степень размытости нечеткого множества. Но здесь важно отметить существенное отличие нечеткой энтропии от классической: нечеткая энтропия содержит нечеткую неопределенность, в то время как классическая представляет сособой случайную вероятность.

Пусть $X=(x_1, x_2, \ldots, x_n)$, $F(X)$ нечеткое множесто на X, $P(X)$ все четкие множества на x,

$R^+ = [0,+\infty)$, for $\forall A = F(X)$,

$\boldsymbol{A(x)}$ – функция принадлежности $\boldsymbol{A}$.

$A^c = A^c(x) = 1 - A(x)$ - дополнение $\boldsymbol{A}$, пусть $[a] = [a](x) = a$ ($a \in [0,1]$) будет постоянное нечеткое множество на X.

Степень нечеткости четкого множества равна 0, потому что его элемент может либо принадлежать, либо не принадлежать этому множеству. Степень нечеткости нечеткого множества [0.5] достигает максимального значения поскольку в этом случае элемент с равной вероятностью может принадлежать множеству или не принадлежать. Очевидно, дополнение такого нечеткого множества А будет иметь ту же степень нечеткости.

Также степень нечеткости подмножества А должна быть монотонной: чем ближе А к интервалу [0.5], тем выше эта степень; чем далее от этого интервала, тем ниже. Основываясь на приведенном выше анализе, предлагается определение нечеткой энтропии:

1. $\forall D \in P(x), e(D) = 0$
2. $e([0,5]) = \max_{A \in F} e(A)$
3. $\forall A, A^c \in F(x)\, if\ A^c is\ the\ sharpened\ version\ of\ A$
4. $\forall A \in F(x), e(A^c) = e(A)$

(14.322)

Так как пересечение А и А' не нулевое, Ягер предложил следующую формулу для нечеткой энтропии:

$$l_y(A) = 1 - \frac{l_p(A, A^c)}{n^{1/p}},$$

$$where\ \forall A, B \in F(x),\ p \geq 1, l_p\ is\ depend\ as \qquad (14.323)$$

$$l_p(A, B) = [\sum_{i=1}^{n} |\ A(x_i) - B(x_i)\ |^p]^{1/p}$$

Аксиомы нечеткой энтропии приняты научным сообществом и стали важным критерием определения любой новой нечеткой энтропии. Сейчас существует множество формул энтропии, выбор конкретной зависит от решаемой задачи.

14.16. Исспользование энтропийных и информационных характеристик при исследовании систем

Состояние любой системы характеризуется наличием различных по своей природе и проявлению случайных воздействий, порождающих неопределенность значений ее параметров. С другой стороны, даже для системы, находящейся в стационарном детерминированном режиме, с позиции стороннего наблюдателя, проявление тех или иных значений ее параметров может трактоваться как стационарный случайный процесс, также характеризующийся определенными значениями их состояний неопределенности. Эти обстоятельства обусловливают возможность и целесообразность применения вероятностных и нечетких подходов к исследованию систем и принятию решений по управлению ими. Сказанное относится к проблемам исследования не только технических систем, но и других типов систем различной природы: социальных, экономических, экологических, биологических и др.

Одним из перспективных направлений исследований различных систем является подход, основанный на анализе изменений энтропии параметров и динамики информационных процессов, порождаемых изменением состояний системы. В настоящее время ввиду актуальности этого направления исследований ведутся интенсивные поиски и разработки таких подходов. Однако внедрение существующих теоретических разработок сдерживается отсутствием удобных вычислительных алгоритмов и методик по их применению на практике. Преодоление этого препятствия позволит получить новый инструмент для исследования систем.

Ниже рассматриваются идеи и методы исследования систем на основе энтропийных моделей и с использованием методов теории энтропийных потенциалов, основное содержание которых изложено в работах В.Л. Лазарева. Такой подход имеет перспективы дальнейшего развития и практического приложения.

Постановка задачи. Основные положения и определения

В предлагаемом подходе моделирование системы по рассматриваемому параметру x осуществляется с помощью отображения множества его варьируемых значений X в неслучайную функцию комплексного энтропийного потенциала (КЭП) L_Δ . Вариации значений параметра могут иметь место как в пространстве, так и во времени. При этом для стороннего наблюдателя это множество может являться случайным. Выражение для L_Δ имеет вид:

$$L_\Delta = \frac{\Delta_e}{X_n} = \frac{K_e \sigma}{X_n} \tag{14.324}$$

где Δ_e — энтропийный потенциал параметра, определяемый как половина диапазона равномерного распределения ($2\Delta_e$) с одинаковой плотностью распределения вероятности $p = 1/(2\Delta_e)$ в этом диапазоне, имеющего такую же вероятностную или информационную энтропию $H(x)$, что и закон распределения рассматриваемого параметра (не следует путать с термодинамической энтропией, которая характеризует направленность процессов теплообмена). X_n — базовое значение параметра, на фоне которого рассматривается состояние системы. В качестве базового значения могут быть выбраны математическое ожидание параметра, диапазон его изменения, предельное или какое-либо номинальное значение. При таком определении величина энтропийного потенциала является обобщенной и унифицированной, на базе закона равномерной плотности, характеристикой состояния неопределенности параметра с любым другим законом распределения. Ее размерность совпадает с размерностью параметра, и она однозначно выражается через его энтропию $H(x)$ следующим образом:

$$\Delta_e = \frac{1}{2}e^{H(x)} = \frac{1}{2}e^{-\int\limits_{-\infty}^{\infty} p(x)\ln p(x)dx} \qquad (14.325)$$

где $p(x)$ — плотность распределения вероятности (закон распределения) значений параметра.

С другой стороны, величина e может быть выражена через вероятностные характеристики параметра в виде

$$\Delta_e = K_e\sigma, \qquad (14.326)$$

где K_e — энтропийный коэффициент, характеризующий дестабилизирующие свойства закона распределения параметра, неопределенность в предсказуемости его значений в зависимости от вида этого закона, $0 \le K_e \le 2.066$. Максимальное значение $K_e = 2.066$ имеет нормальный закон распределения, характеризующийся „наихудшей предсказуемостью" проявлений тех или иных значений параметра относительно других законов. Рекомендации и методики определения значения K_e для различных ситуаций, законов распределений и их композиций приведены в ряде работ. Там же приведены численные значения K_e для ряда законов. σ — величина среднего квадратического отклонения параметра (СКО).

Возможность выражения величины Δ_e через вышеупомянутые характеристики рассеяния является одним из достоинств ее практического использования при описании состояний неопределенности. Для вычисления этой величины зачастую требуется гораздо меньший объем экспериментальных данных и вычислительных ресурсов, чем для вычисления соответствующего значения энтропии. Минимальный объем данных необходим для вычисления СКО, а значения K_e во многих случаях могут быть определены аналитически, исходя из физического смысла, на основании аналогий с подобными явлениями идр.

Частным аналогом понятия энтропийного потенциала является понятие энтропийного значения погрешности. Увеличение Δ_e свидетельствует о возрастании уровня состояния неопределенности параметра, и наоборот.

В отличие от величины Δ_e величина L_Δ является безразмерной, что позволяет использовать ее в качестве критерия энтропийного подобия при сравнении состояний неопределенности различных систем.

Следует отметить, что если X_n = const, то величина L_Δ будет являться масштабным изображением энтропийного потенциала. Если принять также K_e = const, что соответствует пренебрежению изменениями законов распределения параметров на различных этапах существования системы и аппроксимации их каким-либо одним законом (например, законом равномерной плотности с K_e = 1.73), то L_Δ выродится в масштабное изображение величины СКО. В этом случае технологии моделирования и исследования эволюции систем на основе методов теории энтропийных потенциалов сведутся к классическим моделям и технологиям, основанным на использовании методов дисперсионного анализа.

Увеличение модуля КЭП, $|L_\Delta|$, свидетельствует о возрастании уровня состояния неопределенности параметра и наоборот (L_Δ может принимать отрицательные значения в случаях, когда базовое значение X_n является отрицательным, так как величины K_e и σ являются неотрицательными).

Проиллюстрируем изложенный подход следующими примерами.

Рассмотрим вращение точки N на плоскости. В первом случае будем считать, что R — расстояние от оси вращения до точки N и ω — угловая скорость — постоянны. Комплексная модель этого динамического процесса во времени t, будет иметь вид $z=R\exp(j \cdot \omega t)$, а векторная — $y = R\sin(\omega t)$. Соответствующая этому процессу фазовая траектория на фазовой плоскости в системе координат x, $y = dx/dt$ будет представлять собой эллипс (кривая *1* на рис. 14.1).

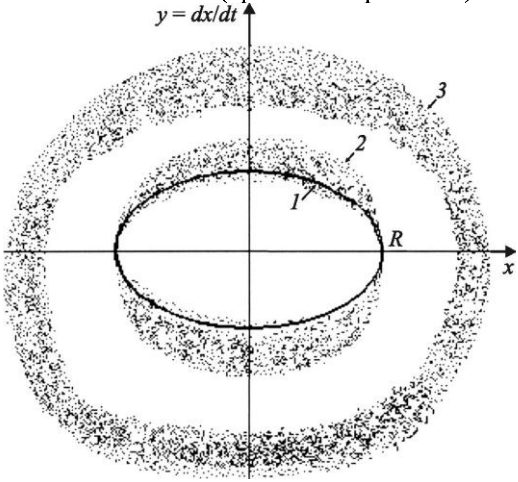


Рис. 14.1. Фазовый портрет системы: *1* — детерминированный режим вращения; *2* — режим вращения с одной нечеткой переменной — ω; *3* — режим вращения с двумя нечеткими переменными — ω и R.

Это же движение с позиции стороннего наблюдателя, не имеющего аналитического описания этого процесса, можно представить совокупностью проекций точки N на какую-либо ось (например, X), являющихся случайными величинами с соответствующим уровнем непредсказуемости или неопределенности. Рассеяние этих величин будет характеризоваться арксинусоидальным законом распределения с K_e=1.11 и $\sigma=R/\sqrt{2}$. Этой ситуации соответствует точка *1* на энтропийной плоскости (в системе координат $\sigma - K_e$) на рис. 14.2.

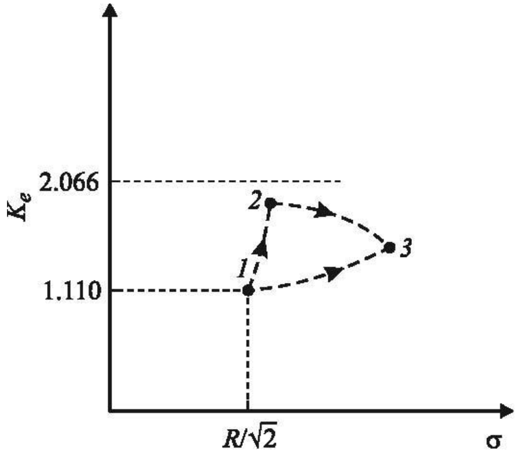


Рис. 14.2. Энтропийный портрет системы.

В геометрической интерпретации значение соответствующего энтропийного потенциала Δ_e будет равно площади прямоугольника со сторонами σ и K_e .

Теперь рассмотрим случай когда ω — угловая скорость — является случайной величиной. Такая ситуация встречается довольно часто в различных практических задачах и связана, например, с изменениями нагрузочных моментов на валах двигателей. Подобные задачи возникают при исследовании кавитации на лопастях гребных винтов с регулируемым шагом в условиях интенсивного волнения и др. В качестве точки N может рассматриваться точка на поверхности лопасти, где разрушение поверхности происходит наиболее интенсивно. В этом случае использование вышеупомянутых аналитических моделей для описания такого движения оказывается затруднительным, так как в них появляется нечеткая, или „размытая“, переменная — ω. Соответствующий фазовый портрет также оказывается „размытым“ и представляется в виде полосы переменной ширины с выраженной неявно границей (полоса *2* на рис. 14.1). Причем процедура его построения существующими методами (например, методом изоклин) потребует больших трудозатрат, чем в первоначальном варианте. Применение методов интервальной логики также оказывается не всегда эффективным. Достаточно просто и наглядно эта ситуация будет характеризоваться изменением энтропийного потенциала. Очевидно, что произойдет изменение закона распределения параметра (наиболее вероятно, что возрастет уровень его непредсказуемости, а следовательно, и значение его энтропийного коэффициента) и возможно незначительное изменение значения величины σ . Динамика точки N в таком установившемся случайном режиме будет характеризоваться точкой *2* на энтропийной плоскости (рис. 14.2) с соответствующей величиной энтропийного потециала. Пунктиром показана возможная энтропийная траектория перехода в это состояние из детерминированного режима.

Теперь рассмотрим общий случай, когда ω и R являются случайными величинами, что также характерно для многих практических задач. Наличие двух нечетких переменных еще больше усложняет использование аналитических моделей, а на соответствующем фазовом портрете появится дополнительная составляющая „размытости“, что делает его еще менее информативным (полоса *3* на рис. 14.1). Наглядность энтропийного портрета остается прежней. Нахождение системы в новом динамическом режиме будет характеризоваться новыми значениями Δ^e, K_e и σ, чему соответствует точка *3* на рис. 14.2. Пунктирными линиями показаны возможные траектории перехода в это состояние из предыдущих режимов.

Переход к величине комплексного энтропийного потенциала L позволяет получить безразмерную характеристику динамического процесса. При этом различные явления могут оказаться подобными. Так, например, динамика какой-либо планеты, вращающейся вокруг своей звезды, по величине L_Δ может оказаться сравнимой с динамикой электрона на орбите какого-либо атома. Таким же образом можно исследовать изменения движений этих объектов вследствие различных возмущений: столкновения планеты с астероидом или изменение орбиты электрона при получении квантов энергии.

Следует отметить, что с использованием изложенного подхода можно исследовать не только динамику вращательного или колебательного движения. Используя понятия величин энтропийных потенциалов, можно исследовать динамику любых объектов, например, газовых и жидкостных потоков. В этом случае необходимо рассматривать проекции перемещений какой-либо частицы, находящейся в потоке, на оси, лежащие в плоскости, перпендикулярной направлению потока. Можно анализировать перемещение и в продольном направлении, только в этом случае необходимо рассматривать проекции скоростей на это направление или проекцию хаотической составляющей поступательного движения. При таком подходе увеличение модуля величины L_Δ будет характеризовать тенденцию возрастания уровня турбулентности потока, и наоборот. Здесь величина комплексного энтропийного потенциала может являться альтернативой критерия Рейнольдса (Re).

Моделирование и исследование систем в пространстве энтропийных потенциалов. Особенности приложений на практике

Состояние неопределенности каждого из рассматриваемых параметров системы, обусловленное особенностями ее существования, будет характеризоваться положением изображающей точки в трехмерной декартовой системе координат: $\sigma - K_e - X_n$. Такую систему координат можно рассматривать как частный случай фазового пространства. Отображение множеств случайных величин во множество точек этого пространства является сюръекцией, так как различные множества случайных величин с различными законами распределений могут иметь одинаковый энтропийный потенциал.

Условия $L_\Delta = C_j$ = const ($j \in J$), соответствующие различным уровням состояний неопределенности системы, разбивают исходное множество точек трехмерного пространства на классы непересекающихся подмножеств точек $M^{(j)}_{(\sigma,K,X)}$, лежащих на одной поверхности постоянного комплексного энтропийного потенциала — изотропной поверхности. Доказательство этого утверждения оформим в виде следующей теоремы.

Теорема 1. Изотропные поверхности не имеют точек пересечения.

Доказательство проведем методом от противного. Предположим, что какие-либо две изотропные поверхности $L_{\Delta 1}$ = const и $L_{\Delta 2}$ = const ($L_{\Delta 1} \neq L_{\Delta 2}$) пересекаются в какой-либо точке K. Тогда в этой ситуации имеет место равенство $L_{\Delta 1}=L_{\Delta 2}$. Получили противоречие. Что и требовалось доказать.

Очевидно, что перемещение изображающей точки по изотропной поверхности, соответствующее изменению состояния системы по причине ее эволюции под действием каких-либо факторов или реализации какого-либо управления, свидетельствует о неизменности степени ее состояния неопределенности в комплексе. Уменьшению степени состояния неопределенности объекта или системы будет соответствовать переход изображающей точки на другую изотропную поверхность с меньшим значением L_{Δ}, и наоборот. Все множество величин комплексных энтропийных потенциалов $M(\sigma,K,X)$ может быть представлено как объединение всех классов подмножеств точек изотропных поверхностей

$$M_{(\sigma,K,X)} = \bigcup_{j \in J} M^{(j)}{}_{(\sigma,K,X)}. \qquad (14.327)$$

Условие X_n = const в геометрической интерпретации соответствует сечению семейства изотропных поверхностей соответствующей плоскостью. В результате образуется семейство кривых — изотроп — в плоской системе координат $\sigma - K_e$, соответствующих условию

$$\Delta_{ej} = C_j = \text{const}.$$

Используя понятия энтропийных потенциалов, можно количественно охарактеризовать изменение состояния системы по рассматриваемому параметру x на каждом отдельном этапе эволюции по ее „информационному следу“. Вывод и доказательство этого положения оформим в виде следующей теоремы.

Теорема 2. Пусть $L_{\Delta 1}$ и Δ_{e1} — исходные величины энтропийных потенциалов системы в начале рассматриваемого этапа; $L_{\Delta 2}$ и Δ_{e2} — конечные величины энтропийных потенциалов системы в конце этапа. Тогда количество информации I, порожденное изменением состояния неопределенности системы на данном этапе, инвариантно относительно соответствующих базовых значений параметра X_{n1}и X_{n2} и равно $I = \ln(\Delta_{e1}/ \Delta_{e2})$.

Доказательство. Определим приращение величины комплексного энтропийного потенциала системы на данном этапе и выразим величины Δ_{e1} и Δ_{e2}через соответствующие им значения энтропий $H_1(x)$

и $H_2(x)$ из выражения (14.325). В результате получим

$$L_{\Delta 1} - L_{\Delta 2} = \frac{\Delta_{e1}}{X_{n1}} - \frac{\Delta_{e2}}{X_{n2}} = \frac{\Delta_{e1} X_{n2} - \Delta_{e2} X_{n1}}{X_{n1} X_{n2}} =$$

$$= \frac{\Delta_{e2} X_{n1}}{X_{n1} X_{n2}} \left(\frac{\Delta_{e1} X_{n2}}{\Delta_{e2} X_{n1}} - 1 \right) = L_{\Delta 2} \left(\frac{\Delta_{e1} X_{n2}}{\Delta_{e2} X_{n1}} - 1 \right) = \quad (14.328)$$

$$= L_{\Delta 2} \left(\frac{X_{n2}}{X_{n1}} e^{I} - 1 \right).$$

В выражении (14.328) величина $I = H_1(x) - H_2(x)$ является мерой количества информации, порожденной изменением состояния неопределенности системы, ее „информационным следом“ на данном этапе. Из выражения (14.328) следует

$$L_{\Delta 1} - L_{\Delta 2} = L_{\Delta 2} \left(\frac{X_{n2}}{X_{n1}} e^{I} - 1 \right),$$

или

$$\frac{L_{\Delta 1}}{L_{\Delta 2}} = \frac{X_{n2}}{X_{n1}} e^{I},$$

или

$$= \frac{\Delta_{e1} X_{n2}}{X_{n1} \Delta_{e2}} = \frac{X_{n2}}{X_{n1}} e^{I}.$$

Откуда следует

$$I = \ln \frac{\Delta_{e1}}{\Delta_{e2}}. \quad (14.329)$$

Что и требовалось доказать.

Полученный результат (14.329) можно развить для удобства практического использования, выразив величины энтропийных потенциалов через соответствующие характеристики рассеяния в соответствии с выражением (14.326)

$$I = \ln \frac{\Delta_{e1}}{\Delta_{e2}} = \ln \frac{K_{e1\sigma 1}}{K_{e2\sigma 2}} = \ln k_{ke} + \ln \frac{\sigma 1}{\sigma 2}, \quad (14.330)$$

где $k_{ke} = \dfrac{K_{e1}}{K_{e2}}$ — коэффициент преобразования закона распределения параметра. Коэффициент k_{ke} является одним из понятий теории энтропийных потенциалов. Значения k_{ke} в ряде случаев могут быть определены теоретически, исходя из физического смысла с использованием аналогий и др., а для некоторых типовых ситуаций могут быть вычислены заранее и табулированы. Строго говоря, величина k_{ke} характеризует статистический режим трансформации закона распределения параметра. Динамика такого процесса будет описываться соответствующим дифференциальным уравнением или передаточной функцией — $W_k(p)$. Поэтому в общем случае величина k_{ke}, например, может быть определена из передаточной функции в виде: $k_{ke}=W_{ke}(p)I_p=0=W_{ke}(0)$. Полученная модель учитывает трансформацию закона распределения параметра в исследуемом процессе, что позволяет повысить эффективность его мониторинга и организации управления.

Изменение величины L_Δ для каждого этапа можно оценить по величине дифференциала комплексного энтропийного потенциала dL_Δ

$$dL_\Delta = \frac{K_e d\sigma}{X_n} + \frac{\sigma dK_e}{X_n} - \frac{K_e \sigma dX_n}{X^2_n} = \frac{\left(K_e d\sigma + \sigma dK_e - \dfrac{K_e \sigma dX_n}{X_n}\right)}{X_n} = \quad (14.331)$$

$$= \frac{K_e d\sigma + \sigma dK_e - K_e \sigma d(\ln X_n)}{X_n} = \frac{(d\Delta_e - \Delta_e d(\ln X_n))}{X_n}.$$

С использованием выражения (14.331) представляется возможным сформулировать положение о неизменности состояния неопределенности системы в результате ее эволюции или реализации какого-либо этапа управления из условия $dL_\Delta=0$. Очевидно, что это условие в геометрической интерпретации описывает перемещение изображающей точки по изотропной поверхности. В результате получим

$$K_e d\sigma + \sigma dK_e - K_e \sigma d(\ln X_n)=0. \quad (14.332)$$

В выражении (14.332) величина дифференциала dK_e характеризует изменение энтропийных или „дестабилизационных“ свойств закона распределения параметра; $d\sigma$ — изменение величины СКО параметра; $d(\ln X_n)$ — изменение величины базового значения параметра в логарифмическом масштабе.

Используя выражение энтропийного потенциала через характеристики рассеяния, то же положение (14.332) можно привести к виду

$$d\Delta e - \Delta e d(\ln X_n) = 0; \quad \frac{d\Delta_e}{\Delta_e} = d(\ln X_n) = \frac{dX_n}{X_n}. \qquad (14.333)$$

Из выражения (14.333) следует, что для нахождения изображающей точки на изотропной поверхности, т. е. для обеспечения постоянного уровня непределенности системы, необходимо, чтобы относительное изменение величины энтропийного потенциала равнялось относительному изменению базового значения.

В ряде задач, когда существуют допускающие дифференцирование аналитические зависимости для входящих в выражение (14.324) величин σ, K_e и X_n от варьируемых параметров системы y_j (j =1,2,...,n), величина dL_Δ может быть определена из выражения

$$dL_\Delta = \frac{\sigma dK_e}{X_n} + \frac{K_e d\sigma}{X_n} - \frac{K_e \sigma dX_n}{X^2{}_n} = \frac{\sigma}{X_n}\sum_{j=1}^{n}\frac{\partial K_e}{\partial y_j}dy_j +$$

$$+\frac{K_e}{X_n}\sum_{j=1}^{n}\frac{\partial \sigma}{\partial y_j}dy_j - \frac{\sigma K_e}{X^2{}_n}\sum_{j=1}^{n}\frac{\partial X_n}{\partial y_j}dy_j. \qquad (14.334)$$

Предложенный подход к исследованию состояний неопределенности обладает высокой „чувствительностью“, так как позволяет уловить тенденцию изменения этих состояний на уровне изменения законов распределений параметров даже при постоянстве их „энергетики“, характеризуемой соответствующими величинами СКО. С другой стороны, повышается наглядность изменений динамики за счет „сжатия“ информации о состояниях системы в стационарных режимах. При этом одна изображающая точка энтропийного портрета будет описывать стационарный процесс сколь угодно большой временной протяженности. Эти свойства энтропийных потенциалов делают их эффективными при оценке состояний и управлении сложными, многофакторными системами в различных режимах.

Рассмотрим пример, иллюстрирующий состоятельность предложенного подхода.

Проведено исследование эволюции экосистемы СанктПетербурга по параметру температурного режима в период с 1865 по 1993 г. В качестве исходной базы были выбраны среднесуточные значения температуры за 10 января по годам указанного периода.

Температурный режим города определяется господствующей „розой ветров“ и является характерным для температурных режимов целого региона.

Для исследования эволюционного процесса исходный период был разбит на три равных временных этапа длительностью 43 года: 1856−1907, 1908−1950 и 1951−1993 гг. Для идентичности условий последующих исследований во всех трех случаях моделирование плотности вероятности распределения температур осуществлялось путем разбиения соответствующих температурных диапазонов на шесть равных интервалов. В качестве базовых значений X_n на каждом этапе использовались средние значения температуры — m_x. При расчетах значений Δ_e для каждого из трех временных этапов вводилась поправка на смещение от недостаточно большого числа наблюдений, попадающих в каждый интервал гистограммы. Результаты исследований представлены в таблице.

Интервалы исследований, гг.	Среднее значение m_x, °C	СКО σ, °C	Энтропийный коэффициент K_e	Энтропийный потенциал Δ_e, °C	Модуль КЭП $\|L_\Delta\|$
1865−1907 1908−1950 1951−1993	−6.96 −8.65 −7.45	6.20 7.36 8.37	1.84 1.82 1.68	11.43 13.42 14.06	1.64 1.55 1.87

Из приведенных результатов следует, что устойчивого повышения температурного уровня (m_x) не наблюдается, и говорить о тенденции тотального потепления за рассматриваемый период преждевременно. С другой стороны, наблюдается отчетливая тенденция повышения уровня нестабильности, непредсказуемости температурных режимов в регионе (столбцы для Δ_e и $|L_\Delta|$). Причина этого обусловлена значительным увеличением разброса температур между отдельными этапами, т. е. „энергетической составляющей“ энтропийных потенциалов (столбец для σ) на фоне менее значительного уменьшения „дестабилизационных свойств“ законов распределения температур (столбец для K_e). Указанная тенденция также диагностируется и по „информационному следу“ системы. В соответствии с выражением (14.329) изменение температурного режима после завершения первого обобщенного этапа характеризуется следующим количеством информации I_{1-2} =ln(Δ_{e1}/ Δ_{e2})=−0.16 (nt), а после завершения второго — I_{2-3}=ln(Δ_{e2}/ Δ_{e3})=−0.046 (nt). Знак „минус“ свидетельствует о нарастании уровня нестабильности; а единица измерения (nt) определяется основанием используемого логарифма. Имеются основания утверждать, что наметилась тенденция к изменению характера климата. Вполне возможно, что наблюдаемые с конца XX в. повышения зимних температур являются следствием вышеуказанных изменений характеристик температурного режима. Более полная диагностика может быть сделана на основании аналогичных исследований соответствующей выборки по температурному режиму за последующий период времени.

Большое практическое значение для применения предложенного подхода к изучению систем является наличие метрологического обеспечения самого процесса исследования. Другими словами, необходимо наличие зависимости для однозначной „точностной" оценки получаемых характеристик исследуемого процесса.

Искомую зависимость можно получить следующим путем. Последовательно прологарифмируем, а затем продифференцируем обе части выражения (14.324). В результате получим

$$\frac{dL_\Delta}{L_\Delta} = \frac{dK_e}{K_e} + \frac{d\sigma}{\sigma} - \frac{dX_n}{X_n}. \qquad (14.335)$$

Все члены выражения (14.335) можно трактовать как величины приведенных погрешностей: γL, γK , $\gamma\sigma$, γX . С учетом существующих обозначений искомая зависимость примет вид

$$\gamma L = \gamma K + \gamma\sigma - \gamma X. \qquad (14.336)$$

Представленный подход к исследованию может быть пролонгирован и на общий случай — многомерную систему. В этой ситуации состояние неопределенности многомерной системы будет характеризоваться m-мерным вектором, координатами которого являются величины комплексных энтропийных потенциалов отдельных параметров L_{Δ_i} (i =1,2,...,m). Так, например, в экосистемах, подобных рассмотренной выше, координатами этого вектора могут являться энтропийные потенциалы температур, уровней грунтовых вод, концентраций тех или иных веществ в атмосфере и грунте, определяемые пространственными или временными факторами. Аналогичным образом можно подойти к исследованию динамики объекта, совершающего сложное перемещение с одновременным многомерным вращением и др. Для организации мониторинга и управления таким объектами и системами необходим единый обобщенный показатель, являющийся функцией этих величин. В качестве такого показателя предлагается использовать величину многомерного комплексного энтропийного потенциала — критерий La_z

$$La_z = \left(\sum_{i=1}^{m} c_i \mid L_{\Delta_i} \mid^{z} \right)^{1/z} = \left(\sum_{i=1}^{m} \left(c_i \left| \frac{\Delta_{ei}}{X_{ni}} \right| \right)^{z} \right)^{1/z} =$$

$$= \left(\sum_{i=1}^{m} \left(c_i \frac{K_{ei}\sigma_i}{\mid X_{ni} \mid} \right)^{z} \right)^{1/z}. \qquad (14.337)$$

где $L_{\Delta i}$ — комплексный энтропийный потенциал i-го параметра; Δ_i — энтропийный потенциал i-го параметра; c_i — весовые коэффициенты, характеризующие значимость, приоритет каждого i-го параметра при описании состояния системы, $c_i \geq 0$ (i =1,2,...,m); z —номер варианта критерия, z =1 или 2.

При z =1 получаем вариант критерия $La_1 = \sum_{i=1}^{m} c_i \mid L_{\Delta_i} \mid$,

При z =2 – вариант $La_2 = \sqrt{\sum_{i=1}^{m} c_i \mid L_{\Delta_i} \mid^{2}}$.

В геометрической интерпретации величина La_2 является модулем или длиной m-мерного вектора, составляющими которого являются значения комплексных энтропийных потенциалов отдельных параметров системы в масштабе их весовых коэффициентов. А величина La_1 является суммой длин модулей этих составляющих, поэтому имеет место условие $La_1 \geq La_2$, причем равенство имеет место в случае, когда m =1. Выбор варианта критерия является прерогативой пользователя. Возрастание величин La_z в обоих вариантах свидетельствует о повышении уровня неопределенности системы, и наоборот. Во всех случаях величина La_z , так же как и значение L_Δ, является безразмерной, что позволяет использовать ее в качестве критерия энтропийного подобия при исследовании состояний неопределенности различных систем. Очевидно также, что при m =1 величина La_z вырождается в модуль значения L_Δ. Разработан и проиллюстрирован подход к моделированию и исследованию различных систем на основе использования понятий энтропийных потенциалов и величины информации, порожденной их изменением. В результате получены наглядные и компактные модели для количественной оценки состояний систем и процессов их эволюции в детерминированных ситуациях и в условиях неопределенности.

15. Основы методологии информационного анализа систем

15.1. Информационный анализ некоторых систем

Возникнув на почве техники связи, теория информации на первом этапе была призвана решать целый ряд выдвинутой этой областью техники практических задач. С помощью теории информации удалось разработать методы построения оптимальных каналов связи, обеспечивающих передачу наибольшего количества информации при наличии помех, методы построения оптимальных кодов, уменьшения содержащейся в сообщениях избыточной информации и т. п. Совокупность этих проблем и методов их решения и составляет фундамент той теории, которую очевидно, можно назвать «классической», в отличие от «неклассической» теории информации, которая формируется в физике, биологии, психологии, лингвистике и других областях.

Основное внимание в дальнейшем изложении будет уделено использованию понятий и методов теории информации именно в «неклассических» областях, которые изучены и разработаны Е.А. Седовым. Однако при этом придется более подробно остановиться и на некоторых «классических» понятиях (например, на понятии «избыточности»), поскольку именно на базе «классической» теории информации возникли и начали развиваться приложения теоретико-информационных идей в упомянутых областях. Примером непосредственного переноса методов теории информации, используемых техникой связи, в область физиологии может служить оценка пропускной способности нервных каналов связи, обслуживающих зрение и слух. С помощью специальных экспериментов установлено, что пропускная способность органов зрения составляет около 10^6 бит/сек, органов слуха — около 10^4 бит/сек. Проведенные с привлеченном методов теории информации психологические опыты показали, что количество осознанно воспринимаемой информации не превышает 50 бит/сек. На языке «классической» теории информации данное свойство сознания может быть сформулировано так: пропускная способность каналов снязи, в которых осуществляется осознанная обработка информации, не превышает 50 бит/сек. Приблизительная оценка информационной емкости головного мозга, произведенная на основе подсчета количества нервных клеток коры больших полушарий (нейронов), показала, что мозг способен хранить количество информации, выражаемое числом порядка 10^{20} бит.

Принципиально новым понятием, введеным теорией информации в современную науку, является понятие *избыточной информации.* К этому понятию нам придется неоднократно возвращаться в последующих разделах, поэтому для уяснения его сущности рассмотрим предложенные теорией информации количественные критерии избыточности и способы ее уменьшения.

Избыточной информацией является та информация, которую можно заранее предсказать. Важно подчеркнуть, что предсказания эти не являются, как правило, однозначными, а даются в вероятностной форме: для русского письменного языка вероятность того, что следующим сообщением будет буква «О», состляет 0,2, вероятность того, что после сочетания «ий» появится интервал, составляет 0,7 и т. п.

Возможность предсказания последующих сообщений (букв) при передаче текста обусловлена наличием *корреляции* (взаимосвязи) между буквами, слогами, словами. В свою очередь, корреляция является результатом действия грамматических и фонетических правил. Именно благодаря корреляции можно заранее предсказать, что в любом русском тексте сочетание «ться» будет появляться очень часто, что после сочетаний «ский» с вероятностью, равной 1 (т. е. с достоверностью), появится интервал, что вероятность сочетания «ткрп» равна нулю, и т. п.

Выявление избыточной информации является экспериментальной задачей, которая решается путем статистического анализа множества сообщений. Так, пока не были исследованы статистические свойства русских письменных текстов, получаемая от каждой принятой (прочитанной) буквы информация наряду с фактически получаемой новой, непредсказуемой информацией $I_Ф$ включала в себя всю невыявленную избыточную информацию $I_и$ и составляла:

$$I_{\max} = I_Ф + I_и = 5 \text{ бит}. \qquad (15.1)$$

Как уже отмечалось, эта величина определялась с помощью функции

$$\sum_i p_i \log p_i,$$

в которой символами p_i обозначались вероятности различных букв (i = А, Б, . . ., Я). При этом считалось, что $p_А = p_Б = \ldots = p_Я$. В соответствии с рассмотренными ранее свойствами функции

$\sum_i p_i \log p_i$ подсчитанная с помощью этой функции величина I была максимальной и составляла для 32-буквенного алфавита в среднем 5 бит на каждую букву. После того как путем статистического анализа множества текстов были установлены реальные значения вероятностей что $p_A = p_Б = \ldots = p_Я$, подсчитанная с помощью той же функции $\sum_i p_i \log p_i$ величина I снизилась до 4 бит. Это произошло потому, что входящая в выражение (15.1) избыточная информация $I_и$ при выявлении реальных значений что $p_A = p_Б = \ldots = p_Я$ была *частично* (в среднем на 1 бит на букву) учтена.

Полное выявление входящей в выражение (15.1) избыточной информации — крайне сложная статистическая задача, требующая учета не только реальных значений что $p_A = p_Б = \ldots = p_Я$, но и взаимной корреляции между буквами, слогами, словами и фразами. Точное значение $I_и$ письменных текстов еще полностью не установлено. По приблизительным оценкам для русских письменных текстов $I_и$ = 4 бита на букву. В соответствии с выражением (15.1):

$$I_Ф = I_{max} - I_и = 1 \text{ бит на букву.}$$

В качестве количественного критерия избыточности теория информации предложила использовать коэффициент избыточности R, определяемый как:

$$R = 1 - \frac{I_Ф}{I_{max}}. \tag{15.2}$$

Значению $I_Ф$ = 1 бит на букву соответствует R = 0,8. Это означает, что информация, содержащаяся в письменных текстах, приблизительно на 80% является избыточной, а потому в принципе предсказуемой путем предварительного учета статистических свойств языка.

Необходимость оценки величин $I_и$ и R была продиктована практическими нуждами техники связи, в которой избыточность сообщений оплачивается лишним временем загрузки каналов связи, лишним расходом мощности и т. п. Для уменьшения этих непроизводительных расходов теорией информации были предложены специальные способы кодирования текстов, позволяющие в той или иной степени уменьшить избыточность передаваемых сообщений. Суть этих способов заключается в замене букв или буквенных сочетаний сериями импульсов, длительность или амплитуда которых тем меньше, чем больше вероятность появления соответствующих букв. В результате такого кодирования резко уменьшаются энергия и время загрузки каналов связи, расходуемые на передачу часто повторяющихся слогов и букв (вопросы кодирования информаци будут подробно рассмотрены отдельно в настоящей работе).

Аналогичные способы уменьшения избыточности применяются при передаче нетекстовых сообщений. Так, например, при передаче сведений о ходе каких-то процессов (изменения скорости, давления, температуры и т. п.) для разгрузки каналов связи оказывается выгодным осуществлять на приемном и передающем конце линии связи автоматическое предсказание (экстраполяцию) хода передаваемых кривых. Тогда для точной передачи данных оказывается достаточным передавать только выявленную на передающем конце разницу между предсказанной и реальной кривой. Тем самым в значительной мере снижается количество информации, загружающей предназначенный для ее передачи канал.

На предсказаниях определенных свойств сообщений основан также способ, получивший название «укрупнение сигнала». Например, вместо часто повторяющегося сочетания четырех букв «ться» можно передавать один кодовый знак. Точно так же заранее можно предвидеть, что в праздничных телеграммах будут часто встречаться слова «поздравляю с наступающим праздником», «желаю здоровья, счастья» и т. п. Заменой этих слов и их сочетаний условными кодовыми знаками удается разгрузить каналы связи и снизить стоимость праздничных телеграмм. Сущность понятия «избыточности» сохраняется неизменной для любых видов информации, хотя форма и способы ее оценки зависят от характера сообщений. Так, например, в телевизионных сообщениях основную часть избыточной информации составляет передача неизменной части телевизионного кадра (например, декораций, на фоне которых происходит действие телеспектакля, и т. п.). Для устранения этой избыточности теория информации предложила осуществлять селекцию неподвижной части изображений и запоминание ее на приемном конце. Тем самым телевизионный канал связи разгружается от избыточной информации и может быть полнее использован для передачи только той информации, которую нельзя заранее предсказать (движения и мимика героев спектакля, события спортивного состязания и т. п.). Данный способ уменьшения избыточности, так же как и все ранее упомянутые способы, основан на предсказании. Чтобы использовать этот способ, необходимо предсказать, что некоторые находящиеся в кадре предметы (пол, потолок, стены, шкаф и т. п.) в течение определенного времени передачи будут оставаться на прежних местах.

При устранении избыточности сообщений необходимо проявлять осторожность. Теория информации показала, что избыточность сообщений может являться не только злом, но и добром. Благодаря избыточности возрастает устойчивость сообщений к воздействиям всякого рода помех. Это свойство можно проиллюстрировать множеством наглядных примеров. Пусть, например, по каналу связи принята такая последовательность букв: ТОМКОЕ. Третья буква переданного слова исказилась из-за помехи. Однако ее значение нетрудно восстановить. Для этого достаточно дождаться появления следующего слова. Если оно окажется БОЛОТОМ, то первое слово следует читать как ТОПКОЕ. То же самое слово может быть прочтено и как ТОНКОЕ, если дальше речь идет, например, о сукне или стекле. Подобное восстановление смысла (которое, кстати сказать, часто осуществляется нами при чтении письменных, а особенно рукописных текстов) возможно благодаря взаимной корреляции букв и слов.

Вообразим искусственный, полностью лишенный избыточности язык. В таком языке *любая* комбинация букв представляет собой осмысленное слово, поэтому изменение одной буквы изменяет его смысл (в реальном языке только около 0,0001% всех возможных буквенных комбинаций образуют осмысленные слова). Тогда было бы почти невозможно выявить опечатки и восстановить первоначальный подлинный текст.

Самым простым и очевидным способом повышения помехоустойчивости сообщений путем увеличения их избыточности является многократное повторение фраз или слов. Этот прием часто используется и в специальных каналах связи, и в обиходной речи при наличии посторонних шумов.

Далее рассмотрим вопрос о том, каким образом методы теории информации, разработанные для нужд техники связи, преломляются и используются в других научных областях.

Тесная связь теории информации с лингвистикой представляется вполне очевидной. Текст — это одна из наиболее универсальных, а потому и наиболее распространеиных форм передачи информации, поэтому задачи создания оптимальных систем передачи сообщений в значительной мере определяются успехами в области статистических исследований языка. В свою очередь, взгляд на язык как на регламентированную определенными правилами стохастичную систему оказывется весьма плодотворным не только для техники связи, но и для самой науки о языке. В настоящее время существуют целые научные школы, базирующиеся на системном подходе к анализу языка.

Сравнительный статистический анализ различных языков породил принципиально новое понятие о *макроструктуре языка.* Оказалось, что, несмотря на множество частных различий в *микроструктуре,* обусловленных действием конкретных грамматических и фонетических правил разных языков, более обобщенные статистические характеристики языков (кривые распределения вероятностей, моменты распределений, корреляционные характеристики и т. п.). весьма сходны между собой. В этом проявляется действие и общность законов более высоких порядков, относящихся к человеку вообще и не зависящих от национальной принадлежности и особенностей конкретного языка. Это касается в первую очередь общих закономерностей, присущих процессам мышления, связям между мышлением и речедвигательными органами и, наконец, законам отражения окружающей действительности сначала в мысли, а затем в языке.

На базе статистического подхода к изучению языковой структуры возникло направление исследований, связанное с построением *моделей языка*. Такие модели строятся на основе определенных формализованных правил (алгоритмов). Они не несут никакой смысловой информации, но сохраняют все формальные признаки реального языка. Эти модели помогают, с одной стороны, более глубоко осознать свойства реальных языков, а с другой — найти пути построения оптимальных искусственных языков, необходимых, в частности, для эффективного использования электронных машин.

Весьма интересны результаты применения методов теории информации в психологии. Так же как и в лингвистике, информационно-статистический анализ в области психологии позволяет перейти от качественных оценок к количественным измерениям и способствует выявлению самых общих закономерностей и свойств. Многие из проведенных психологических исследований посвящены особенностям восприятия, обработки и запоминания информации человеком. Результаты этих исследований представляют большой интерес как для психологов, так и для инженеров, поскольку они, с одной стороны, углубляют научные представления о психологии восприятия, а с другой — открывают возможности создания более совершенных технических устройств. Вполне естественно, что как в том, так и в другом случае основой сравнительных оценок получаемых экспериментальных данных является введенная теорией информации количественная мера. Среди достигнутых в данной области результатов наиболее важными следует считать психологические эксперименты, проведенные В. Хиком и Р. Хайменом. Эти опыты показали, что длительность психологической реакции на внешние сигналы пропорциональна количеству информации, содержащейся в этих сигналах. При этом была экспериментально подтверждена правомочность определения информации как «меры неожиданности» получаемых сообщений, введенного ранее теорией информации на основе общих соображений.

Эксперименты, о которых идет речь, заключались в следующем. Перед испытуемым размещалось световое табло, на котором поочередно зажигались чередующиеся в случайном порядке разноцветные сигналы. Задача испытуемого состояла в том, чтобы нажимать одну из N клавиш после того, как появился некоторый сигнал (клавиши и цветовые сигналы находятся во взаимно-однозначном соответствии). В простейшем случае задавалось равновероятное распределение сигналов. При этом длительность поисков нужных клавиш возрастала пропорционально $\log N$. Как отмечалось выше, в таком же соотношении при равновероятном распределении увеличивается и количество информации (в соответствии с формулой Хартли или с формулой Шеннона при одинаковых p_i).

Данная зависимость длительности психологических реакций от числа равновероятных стимулов была установлена Д. Меркелем еще в конце XIX века. После появления количественной меры информации возникло естественное желание связать длительность психологических реакций с количеством информации, содержащейся в стимулирующих сигналах. Именно эта задача была решена в упомянутых экспериментах В. Хика (для равновероятного распределения) и Р. Хаймена (для различных распределений вероятностей p_i). Во всех случаях длительность психологических реакций испытуемого определялась количеством получаемой информации в соответствии с формулой

$$T = AI = -\mathrm{A}\sum_i p_i \log p_i.$$

Например, длительность реакции уменьшалась в том случае, когда один из сигналов (допустим, красный) возникал намного чаще других. Объясняется это тем, что в ходе эксперимента испытуемый «привыкает» к появлению красного цвета и к нажатию на соответствующую ему клавишу. Сигнал красного цвета перестает быть «неожиданным», испытуемый заранее готов правильно на него реагировать. Правда, вероятность возникновения остальных сигналов при этом сокращается, и сигналы особенно неожиданные (в силу малой их вероятности) могут задерживать реакцию испытуемого на продолжительный срок. Но поскольку такие сигналы редки, затраченное на них время будет невелико. В целом время реакции на сигналы с неравномерным распределением вероятностей будет меньше, чем в случае равновероятного появления тех же сигналов. В том же соотношении уменьшается и количество получаемой испытуемым информации, которая в данном случае характеризует *усредненную по времени неожиданность* стимулирующих сигналов. При равновероятном распределении все сигналы в равной степени неожиданны. При различных значениях вероятностей количество информации уменьшается ввиду того, что некоторые сигналы можно заранее ожидать. Следует подчеркнуть, что установленная таким образом зависимость длительности психологических реакций от количества информации, содержащейся в стимулирующих сигналах, представляет собой не зависящую от воли испытуемого объективную закономерность.

Трудно переоценить значение и плодотворное влияние идей теории информации в развитии самых разнообразных научных областей. Не случайно одно из самых знаменательных научных событий прошлого столетия — расшифровка наследственного генетического кода — совпало по времени с периодом бурного развития теории информации. Ключом к расшифровке генетического кода послужила простая идея о том, что для образования из четырех имеющихся «букв» (т. е. четырех различных нуклеотидов) двадцати «слов», соответствующих двадцати входящим в состав белков аминокислотам, каждое «слово» должно состоять минимум из трех букв. Эксперименты подтвердили это предположение, показав, что для передачи наследственной информации природа и в самом деле использует четвертичный трехзначный код. При таком коде число возможных «слов» (комбинаций из четырех «букв» алфавита по три «буквы» в «слове») составляет $N_t = 4^3 = 64$. Этот «словарь» обладает некоторой избыточностью, поскольку программы включения 20 аминокислот в процессы синтеза обходятся всего 22 «словами» из возможных шестидесяти четырех (в число 22 помимо кодов двадцати аминокислот входят команды начала и окончании считывания матрицы). Но лучше избыточность, чем недостаток: при переходе от трехзначного к двузначному четвертичному коду «слов» оказалось бы меньше, чем необходимо ($N_2 = 4^2 = 16 < 22$). Именно так рассуждал Г. Гамов, впервые выдвинувший гипотезу о принципах построения генетического кода, которую несколько лет спустя экспериментально подтвердил М. Ниренберг.

Нельзя утверждать, что не будь теории информации, не было бы и этих, в сущности довольно простых, рассуждений (до появления теории информации подобные задачи решались комбинаторикой). Однако и рассуждения эти родились не на голом месте, а на глубоко вспаханной и щедро удобренной почве общих информационных идей.

Важна была убежденность в том, что самой простой азбуки достаточно для передачи потомкам тех наследственных «наставлений», согласно которым живет и развивается сложнейший живой организм. Теория информации и кибернетика породили эту научную убежденность. Благодаря им таинственный ген перестал быть магическим необъяснимым термином, а превратился в материальную субстанцию, в объект исследований, содержащий в себе пусть сложный, но поддающийся расшифровке наследственный код. Именно этот информационный подход служил нитью Ариадны, которая вела генетиков и биохимиков сквозь лабиринт гипотез и экспериментов к конечной цели — расшифровке двадцати, пожалуй, самых важных в мире «слов». Путем огромного количества досконально продуманных и с чрезвычайной тщательностью выполненных экспериментов удалось установить, какой именно последовательности из трех нуклеотидов ДНК соответствует включение в процесс синтеза каждой из 20 участвующих в синтезе аминокислот.

Проделанный А.Е. Седовым настоящий краткий обзор приложений теории информации в различных областях науки далеко не исчерпывает всех порожденных этой теорией новых идей. Здесь были рассмотрены лишь главные достижения науки, которым в последующем поступательном движении науки предстоит сыграть наиболее важную роль. Дальнейшее развитие «классической» теории информации и ее «неклассических» приложений безусловно должно принести еще немало самых разных, в том числе и неожиданных, плодов. Вместе с тем развитие теории информации не следует представлять себе как некое триумфальное шествие по гладкой дороге одних лишь побед и почестей.

Имеющийся на сегодняшний день опыт конкретных информационных исследований уже позволил выявить многие методологические трудности и выдвинул ряд новых, пока еще не решенных задач. Трудности многих количественных информационных оценок конкретных объектов связаны со сложностью учета количества информации на всех уровнях организации исследуемых систем. В качестве примера можно привести попытки расчета количества информации в живых организмах с целью оценки степени их усложнения в процессе развития от оплодотворенной яйцеклетки (зиготы) до стадии взрослой особи. При оценке результатов, достигнутых авторами подобных работ, можно вполне согласиться с утверждением М. Аптера о несовершенстве проведенных расчетов: в них налицо произвольность выбора молекулярного уровня, не учтены взаимосвязи между молекулами, группами атомов в хромосомах и т. д.

Тем не менее, саму тенденцию разработки количественных методов в биологии и привлечения меры количества информации для анализа биологических процессов следует безусловно приветствовать, невзирая на вполне понятную неуверенность первых шагов в этом направлении. В этой связи следует заметить, что существующий аппарат теории информации пригоден в принципе для оценки количества информации лишь на нескольких заданных структурных уровнях статистических систем. Так, например, при оценке избыточности письменных текстов учитывается статистическая связь между буквами внутри слова или на «стыках» слов (т. е. правила грамматических согласований и образований фонем). Затем переходят на связи более высоких структурных уровней языка: связь между фразами, между понятиями, определяемыми с помощью миогих фраз, и т. д. Для подобных оценок непригодна принятая теорией информации количественная мера игнорирующая при оценках количества информации содержащиеся в ней ценность и смысл. На первом этапе становления теории информации такой подход был вполне оправданным: именно благодаря абстрагированию от конкретных свойств информации, ее ценности, смысла науке удалось ввести единую меру количества информации, сделать решающий шаг в исследовании энтропийно-информационных процессов, подобно тому как абстрагирование от конкретных свойств механического движение (неизбежности трения) в свое время позволило Галилею сделать в изучении общих законов механики столь же решающий шаг.

Вполне естественно, что главной задачей следующего этапа развития теории информации является переход от обобщающих определений информации к разграничению ее видов и свойств. Точно так же обобщение понятия «энергия» не избавляет науку от необходимости подробных исследований различия ее форм. Именно эту задачу пытаются решать теории семантической и прагматической информации. Первая из названных теорий стремится решить проблему количественной оценки информации с учетом ее смысла, а вторая — с учетом ценности. Применение аппарата этих теорий к анализу процессов развития на различных структурных уровнях, по-видимому, позволит со временем решить ряд немаловажных проблем. Другой методологической трудностью, ограничивающей применение теории информации в исследованиях процессов развития, происходящих в конкретных системах, является отсутствие совершенного математического аппарата, который позволял бы производить количественную оценку приращения информации при взаимодействиях систем (в частности, при взаимодействии системы с внещней средой).

Упомянутая ранее теорема Хинчина говорит лишь о том, что в процессе взаимодействий происходит приращение информации, равное отрицательному приращению энтропии и определяемое как:

$$\Delta I = - \Delta H = H\,(X;\,Y) - H\,(X) - H(Y),$$

где $H(X)$ и $H(Y)$ — значения энтропии множеств X, Y до их взаимодействия, а $H\,(X;\,Y)$ — значение энтропии системы, возникшей в результате взаимодействия исходных систем.

Методологические трудности исследований процессов развития обусловлены тем, что на сегодняшний день разработан недостаточный аппарат, позволяющий определять ΔH для конкретных случаев взаимодействий. На первом этапе развития теория информации и не ставила перед собой подобной задачи, поскольку она рассматривала только каналы связи, в которых количество информации никогда не возрастает, а может лишь уменьшаться за счет воздействия помех. Именно это обстоятельство послужило поводом для скептических выводов некоторых ученых, в частности высказывания К. Уоддингтона о том, что при анализе процессов биологического развития «теория информации как таковая бесполезна..., хотя слово «информация»... с успехом заменяет такие выражения, как «количество разнообразия» и «специфичность».

Рассмотренные нами в предыдущих разделах примеры опровергают этот взгляд. Они показывают, что ряд немаловажных заключений относительно общих свойств процессов развития позволяет сделать именно информационный подход.

От этого подхода можно ожидать гораздо больших успехов; если сегодня теория информации «как таковая» не разработала совершенного аппарата, позволяющего учитывать прирост информации в процессах развития, то завтра она его создаст. Дальнейшим шагом будет разработка методов количественной оценки уменьшения энтропии (и соответствующего увеличения количества информации) в процессах взаимодействия и развития конкретных материальных систем на различных структурных уровнях их организации.

Оценивая влияние теории информации на развитие физики, следует прежде всего отметить углубление представлений о сущности энтропии. Этот вопрос был рассмотрен подробно ранее. К сказанному следует лишь добавить, что появились работы, в которых делаются небезуспешные попытки новой трактовки термодинамики на основе теории информации. Фундаментальное значение для внедрения идей теории информации в физику имеет работа Д. С. Лебедева и Л. Б.

Левитина, посвященная оценке информационного содержания квантов энергии электромагнитного поля. Проведенный в этой работе анализ является подтверждением и дальнейшим развитием высказанной ранее в общем виде Е.А Седовым идеи о том, что процесс излучения — это процесс превращения неупорядоченной, «энтропийной» тепловой энергии в «информационную» энергию электромагнитного поля. Тем самым подтверждается еще и тот факт, что *первопричина феномена информации в природе — порождающая излучения разница температур.*

Вопрос о влиянии теории информации на развитие физики не ограничивается рассмотрением новых методов решения тех или иных более или менее частных проблем. Благодаря внедрению теории информации в физику само слово «физика» снова приобретает некогда вложенный в него чрезвычайно широкий смысл. В переводе с греческого «физика» — это «природоведение». На первом этапе развития физика и в самом деле включала в себя начала целого ряда ставших затем полностью самостоятельными естественных наук. В недрах молодой физики зародились и химия, и астрономия, и целый ряд других отраслей «природоведения», выросших со временем в целые самостоятельные области знаний.

Однако в диалектическом процессе развития научной мысли тенденция к дифференцированию и специализации наук необходимо должна сочетаться с противоположной тенденцией интеграции научных представлений и знаний. Проявлением именно этой тенденции было рождение теории информации, которая с полным правом вторглась и в физические, и в химические, и в астрономические проблемы, оперируя заново переосмысленным всеобъемлющим физическим понятием «энтропии».

Мало этого. Слияние методов физики и теории информации включает в сферу деятельности новой «физики-природоведения» и решение физиологических и психологических проблем. Именно такой путь развития физики предрекал в свое время один из выдающихся ее представителей Вернер Гейзенберг. В 1948 г., когда теория информации была еще в колыбели, в лекции, прочитанной студентам Цюрихской высшей школы, он сказал: «Допускаем ли мы, что кроме атомов существует нечто другое, например душа, или современное атомное учение тоже должно утверждать (подобно Демокриту.— *Е. С.),* что существуют только атомы и пустота?». «Если мы намереваемся описать процессы жизни или духовные процессы, то нам необходимо расширить математический аппарат. Вполне возможно, что это будет сделано путем введения, наряду с прежними, еще и других понятий, которые можно было бы без противоречий с ними связать».

«Другим понятием», объединившим атомные, биологические и даже духовные процессы, как раз и является, по нашему мнению, понятие «информации», которое вводится с помощью статистической количественной меры. Именно поэтому среди всех существующих и развивающихся в настоящее время направлений научных исследований теория информации ближе всего подошла к решению важнейшей научной задачи: «...Исследовать и исследовать, каким образом связывается материя, якобы не ощущающая вовсе, с материей, из тех же атомов (или электронов) составленной и в то же время обладающей ясно выраженной способностью ощущения». Как будет показано ниже, в основу этих исследований может быть положено изучение общих закономерностей эволюционных (развивающихся) процессов, выражающихся в диалектическом переходе случайных связей в связи детерминированные и находящих свое отражение в свойствах функции $\sum_i p_i \log p_i$.

15. 2. Объективная упорядоченность статистической системы и информация, извлеченная из ее наблюдений

Проведенный в предыдущем пункте анализ показывает, что используемая для количественных оценок энтропии и информации функция

$$\sum_i p_i \log p_i$$

является вероятностной мерой более высокого порядка сложности, нежели вероятность. Поэтому можно считать вполне закономерным тот факт, что процесс «осознания» этой введенной в науку Л. Больцманом и использованной в новом качестве К. Шенноном меры в значительной степени повторяет эволюцию взглядов на само понятие вероятности. Долгое время вероятность считалась субъективной мерой неполноты сведений об исследуемых явлениях; лишь постепенно стал осознаваться ее объективный смысл. Аналогичную трансформацию претерпело понятие количества информации, превратившись из субъективной «меры неведения» в объективную меру неравномерности распределения энергии и упорядоченности статистического движения элементов систем.

Совмещение двух разных аспектов (субъективного и объективного) в едином понятии «информация» послужило причиной противоречивых интерпретаций информации и энтропии. Так, например, Л. Бриллюэн предложил новое определение энтропии как «меры недостатка информации о действительной структуре системы». Противоречивость такого толкования энтропии представляется нам очевидной: объективная характеристика системы, каковой является энтропия, ставится в зависимость от такого субъективного фактора, как степень осведомленности наблюдателя, исследующего эту систему.

Не менее определенно высказался по этому поводу Д. Пирс: «Как только наблюдатель выявил что-нибудь в физической системе, так энтропия системы снизилась, ибо для наблюдателя система стала менее неупорядоченной». Пытаясь развить эту концепцию, Я. П. Терлецкий приходит к следующему выводу: «Энтропия, как и всякая величина, определяемая статистически через вероятности, имеет относительный характер. Нельзя спрашивать: «Какова энтропия этой системы?», не оговаривая одновременно тех сведений, которые у нас имеются об этой системе. Если нам априори известно, что система находится в равновесном состоянии, то, интересуясь ее энтропией, мы имеем в виду ее равновесную энтропию. Если же о системе у нас имеются дополнительные сведения, то можно интересоваться неравновесной энтропией системы, при этом существенно то, что потеря (забывание) этих сведений или просто их игнорирование приводит к иному (равновесному) значению энтропии. И в этом нет ничего удивительного, так как вообще в теории вероятностей вероятности мгновенно переоцениваются, как только мы пожелаем использовать дополнительные сведения, полученные из наблюдений». Приведенные утверждения Я. П. Терлецкого не вызывали бы возражений, если бы речь шла не об объективном значении энтропии, а лишь о наших оценках, безусловно зависящих от имеющихся у нас сведений о системе. В самом деле, как мы уже убедились, наши оценки энтропии зависят от наших априорных сведений, и если заранее ничего не известно, мы вынуждены считать, что все вероятности одинаковы и потому энтропия равна максимальной величине. Однако помимо наших расчетов существует не зависящее от них положение вещей; фактическое значение энтропии материальных систем есть частица объективной истины — один из физических параметров, не зависящих от нас.

Между тем Я. П. Терлецкий не усматривает никакой разницы между энтропией расчетной (разумеется, зависящей от нашей осведомленности) и энтропией «фактической» (которой до нашей осведомленности и наших расчетов никакого дела, разумеется, нет). Вывод, к которому приходит этот автор, формулируется им так: «Это скачкообразное изменение таких, казалось бы, чисто физических величин, как энтропия, только в результате изменения нашего желания использовать дополнительные сведения происходит вследствие того, что любая статистическая теория есть лишь особый, вероятностный способ познания объективных закономерностей природы, т. е. особый статистический способ отражения в нашем сознании возможных объективных, протекающих независимо от нашего сознания физических процессов».

Вывод малоутешительный. Что толку от этого «особого способа познания», «особого способа отражения в нашем сознании объективных... физических процессов», если в результате их применения такой важный физический параметр, как энтропия, перестает быть объективной величиной? При такой трактовке все встает «с ног на голову»: теперь уже не объект дает наблюдателю сведения о своей энтропии, а наблюдатель по собственному усмотрению «оделяет энтропией» объект.

Взгляды Л. Бриллюэна, Д. Пирса и Я. П. Терлецкого имеют один общий исток. Все названные авторы исходят из того, что информация неразрывно связана с энтропией. И в этом все трое безусловно правы. Однако одновременно с этим все они находятся в плену традиционного обиходного, а потому существенно суженного представления об информации как о сведениях, которыми располагает некоторый субъект. А раз энтропия связана с информацией, а информация — это осведомленность того или иного субъекта, значит, и энтропия находится в связи со сведениями, которыми располагает или которые получает субъект. Таким образом, сужение понятия «информация» до «традиционного» толкования информации как меры приобретенных кем-то знаний в конечном счете приводит к описанному методологически неубедительному толкованию физической энтропии.

Однако развитие теории информации и ее многочисленных приложений показало, что информация является не только «мерой неведения», «мерой неожиданности» или каких-либо других чисто субъективных свойсти. Как было показано нами ранее, ***информация является еще и мерой упорядоченности движения, происходящего в объективно протекающих процессах***. Стоит лишь взять за основу данное понимание информации, и физическая энтропия перестает зависеть от наших желаний и знаний, а связь энтропии и информации приобретает не зависящий от субъекта, вполне объективный смысл. При этом и самого субъекта, изучающего энтропию какой-то системы, нужно рассматривать как систему. От полученной при наблюдениях информации в самом деле ***уменьшается энтропия***, по не ***энтропия объекта исследований, а энтропия представлений о нем***.

Первые попытки разграничения объективного и субъективного аспектов энтропийно-информационных соотношений были предприняты Б.В. Бирюковым, А.Д. Урсулом, В.М. Глушковым, Е.А. Седовым. По мере дальнейшего развития теории информации и внедрения ее методов в различные области науки разграничение объективного и субъективного аспектов информации перестает быть вопросом абстрактной теории и превращается в практический вопрос. Подтверждением этого может служить следующее высказывание У. Филдса и У. Эббота о затруднениях, возникающих при попытках количественной оценки информации биологических систем: «Было введено представление о меняющемся объеме информации, причем количество информации, полученной в результате наблюдений ($\Delta I_н$), было определено как величина, обратная дисперсии результатов. На первый взгляд это представление противоречит идее о высокой информационной ценности маловероятных данных. Однако это вовсе не так, потому что при подсчете полного количества информации, полученной при наблюдениях, результаты которых описываются кривой распределения с острым пиком, всегда учитываются и лежащие в крайних точках кривой маловероятные наблюдения, имеющие высокую информационную ценность».

Данное утверждение содержит по меньшей мере две принципиальные неточности:

1. Если под «полным количеством информации» понимать энтропию всех сообщений о результатах исследования системы, то энтропия (т. е. полное количество информации) действительно будет уменьшаться по мере уменьшения дисперсии кривой распределения вероятностей, получаемых при наблюдении данных. Так, например, в случае нормального закона распределения вероятностей уменьшение энтропии (которая согласно приведенной выше формулировке У. Филдса и У. Эббота и является полным количеством информации $\Delta I_н$) при уменьшении дисперсии от величины σ_1 до величины σ_2 согласно выражению (13.97), определится как:

$$\Delta I_н = -\Delta H = \ln \sigma_1/\sigma_2 \quad (\text{при } \sigma_1 > \sigma_2).$$

Следовательно, несмотря на наличие «маловероятных наблюдений, имеющих высокую информационную ценность», определяемое таким образом полное количество информации для кривой вероятностей «с острым пиком» (т. е. с малой дисперсией) будет меньше, чем для плавной кривой.

2. Исходя из приведенных рассуждений У. Филдса и У. Эббота, можно прийти к следующему парадоксальному выводу: наибольшую

информационную ценность для наблюдателя представляет не установление главных закономерностей, присущих объекту, которым соответствует «пик» вероятностного распределения, а случайные отклонения, «лежащие в крайних точках кривой». Как будет показано ниже, парадокс этот возник оттого, что правильно интуитивно найденная обратно пропорциональная зависимость количества информации от дисперсии результатов получила в этом рассуждении ошибочную трактовку.

Исходным моментом отмеченных противоречий является нечеткое разграничение двух аспектов в оценке количества информации в процессе наблюдений статистических систем. *Дело в том, что в процессе наблюдений взаимодействуют два вида информации: информация, объективно существующая в наблюдаемой системе (упорядоченность ее структуры), и информация, которую получает в процессе исследований наблюдающий систему субъект*.

Взаимосвязь количества информации, объективно сохраняемой данной системой, с количеством информации, получаемой при ее наблюдении, была впервые исследована Е.А.Седовым. Покажем, что, опираясь на основные выводы работы Е.А.Седова, можно устранить противоречия, содержащиеся в приведенных нише высказываниях У. Филдса и У. Эббота, уточнив тем самым методологию применения количественной меры информации в исследованиях биологических систем.

Согласно результатам работы Е.А.Седова количество информации, определяемой как избыточная (с точки зрения эффективного использования пропускной способности информационных каналов), одновременно является объективной мерой упорядоченности движения элементов и структурной целостности систем. В самом деле, избыточностью текста, составляющей в среднем около четырех бит на букву, оплачивается грамматическая и фонетическая целостность языка. Ту же самую мысль выеказывал Г. Ферстер, предлагавший использовать ***коэффициент избыточности в качестве меры организованности систем***.

До тех пор пока в результате статистических исследований системы не выявлено полное количестно содержащейся в ней избыточной информации $I_и$, это количество характеризует максимальную *неопределенность представлений* наблюдателя об упорядоченности данной системы $H_{н\,max}$. Таким образом:

$$H_{н\,max} = I_и. \tag{15.3}$$

Так, например, до тех пор пока на основе исследований множества русских текстов не было установлено, что

$p_А = 0{,}075$; $p_Б = 0{,}017$; $p_В = 0{,}046$ и т. д.,
считалось, что усредненная неопределенность появления каждой буквы составляет:

$$H_{\max} = -\sum_{i=1}^{i=32} \frac{1}{32} \log_2 \frac{1}{32} = 5 \text{ бит} \qquad (15.4)$$

(при этом принималось $p_А = p_Б = \ldots = p_Я = \frac{1}{N}; N = 32$). Потом оказалось, что около четырех бит (из пяти) предсказуемы; следовательно, неопределенность представлений наблюдателя, составлявшая в начале исследований $H_{н\ \max}$=4 бита на букву, в результате проведенных статистических исследований была устранена. Другими словами, 4 бита на букву — это та информация об упорядоченной статистической структуре системы, которая была приобретена в результате исследований статистических свойств языка (эту приобретенную в результате статистических исследований информацию мы в дальнейшем будем называть «информацией наблюдений» и обозначать через $\Delta I_н$ (см., например, соотношения (15.5), (15.6) и т. д.).

Оставшийся 1 бит на букву — это непредсказуемая информация, обусловленная объективной стохастичностью языка *($H_о$)*.

Таким образом, величина

$$H_о = H_{\max} - H_{н\ \max}$$

является ***количественной мерой объективных стохастических свойств***.

Из всего сказанного следует, что количество информации, получаемой в процессе наблюдений, зависит как от фактической неопределенности (объективной стохастичности) наблюдаемой системы $H_о$, так и от неопределенности исходных представлений о статистических свойствах системы $H_н$. Поскольку в процессе наблюдений получаемая информация $I_н$ уменьшается вместе с уменьшением неопределенности представлений наблюдателя $H_н$, $I_н$ и $H_н$ естественно считать функциями времени *(t)*. Благодаря свойству аддитивности информации и энтропии связь между $I_н(t)$, $H_н(t)$ и $H_о$ будет выражаться простым соотношением

$$I_н(t) = H_о + H_н(t). \qquad (15.5)$$

Если в начальный момент наблюдений t_0 отсутствуют какие-либо предварительные сведения о статистических свойствах исследуемой системы, то все исходы имеют для наблюдателя равную вероятность, поэтому:

$$I_н(t_0) = H_{max} = H_0 + H_{н\,max}(t_0). \tag{15.6}$$

Примером ситуации, соответствующей условию (15.6), может служить начальная стадия информационно-статистического анализа письменных текстов, когда реальные вероятиости появления различных букв были еще не известны. Появление каждой из 32 букв алфавита имело для наблюдателя равную вероятность и несло количество информации $I_н(t_0) = H_{max}$, причем H_{max} определяется выражением (15.4).

По мере получения, в процессе наблюдений, статистических данных субъективная неопределенность представлений наблюдателя $H_н(t)$ уменьшается, стремясь в пределе к нулю. Согласно выражению (15.5) условию $H_н(t)\to 0$ соответствует условие $I_н(t) \to H_о$ (здесь и далее знак $\to$ означает, что значения соответствующей функции стремятся к некоторому пределу (или бесконечности)).

Обозначив через $t_к$ конечный момент наблюдений, в который полностью устранена первоначальная неопределенность представлений наблюдателя о статистических свойствах системы, можно записать: $I_н(t_к) = H_о$. Процесс перехода от $I_н(t_0)=H_{max}$ к $I_н(t_к)=H_о$ может быть проиллюстрирован тем же примером статистических исследований письменных текстов. После того как в результате статистического анализа множества текстов были установлены фактические значения вероятностей появления всех букв, количество информации $I_н(t_1)$, получаемой наблюдателем при появлении каждой буквы, уменьшилось и стало составлять не 5 бит, а около четырех. Произошло это за счет того, что входящая и выражение (15.5) субъективная неопределенность представлений о статистических свойствах текстов $I_н(t_1)$ в результате установления реальных значений вероятностей $p_А$, $p_Б$, …, $p_Я$ была частично устранена (уменьшилась примерно на 1 бит). Дальнейший статистический анализ текстов позволил учесть взаимную корреляцию букв в пределах четырехбуквенных сочетаний. При этом произошло дальнейшее уменьшение неопределенности представлений до $H_н(t_2)$, и величина $I_н(t_2)$ стала составлять около трех бит на букву.

Приблизительная оценка корреляции между всеми буквами и словами письменных текстов показала, что при полном учете корреляции (т. е. при полном устранении неопределенности представлений наблюдателя $H_н$)

$$I_н(t_к) = H_0 \cong 1 \text{ бит на букву.}$$

Таким образом, после устранения неопределенности представлений

$H_н$ ($t_к$) количество информации, получаемой в процессе дальнейшего наблюдения, соответствует объективной стохастичности исследуемой системы $H_о$. Поскольку объективная стохастичность наблюдателю уже известна, количество *новых* сведений, получаемых при дальнейших наблюдениях, будет равно нулю. Данную ситуацию можно описать так:

$$I_н(t_к) = I(t_{к+j}); \qquad \Delta I_н(t_{к+j} - t_к) = 0,$$

где j = 1, 2, 3...

Так, например, фиксация и статистический учет скоростей огромного количества молекул газа, помещенного в замкнутый объем, с постоянным давлением и температурой, не дает наблюдателю никакой дополнительной информации, ввиду того что полученная им статистическая кривая совпадает с кривой Максвелла (предполагается, что кривая Максвелла с полной достоверностью отражает объективную закономерность и что она была известна еще до начала наблюдений).

Таким образом, количество информации, получаемой наблюдателем, зависит в конечном счете от неопределенности его первоначальных представлений о характере наблюдаемого явления, т.е.

$$\Delta I_{н\,\max}(t_к - t_о) = I_н(t_о) - I_н(t_к) = H_{\max} - H_о = H_{н\,\max}. \tag{15.7}$$

С другой стороны, ***разность*** H_{max}—$H_о$ может служить ***статистической мерой*** существующего в наблюдаемой системе ***порядка*** ($H_н$). Этот порядок определяется количеством объективно существующей информации $I_о$, сохраняемой данной системой:

$$I_о = H_п = H_{\max} - H_о. \tag{15.8}$$

Сопоставляя (15.7) и (15.8), получаем:

$$\Delta I_{н\,\max}(t_к - t_о) = H_п = I_о. \tag{15.9}$$

Полученное соотношение говорит о том, что ***предельное количество информации****, которое может быть получено в результате статистических исследований той или иной системы,* ***равно количеству информации, сохраняемой данной системой***.

Если система не обладает объективной упорядоченностью движения (т. е. не сохраняет информации), первоначальное представление о системе (т. е. предположение о том, что $p_1 = p_2 = ... p_N$) совпадает с объективной неопределенностью системы. В этом случае в выражении (15.5) $H_н(t_0)$=0, поэтому

$$I_н(t_о) = H_о = H_{\max}, \tag{15.10}$$

т. е. количество информации, получаемой на начальной стадии наблюдений t_0, соответствует объективной неопределенности системы и в процессе последующих наблюдений сохраняется неизменным:

$$I_{н}(t_0) = I_{п}(t_{к}) = H_{\max} = H_0. \tag{15.11}$$

Наглядной иллюстрацией взаимосвязи объективного состояния системы и информации, получаемой при ее наблюдении, может служить процесс растворения, работа которого выражается известным соотношением:

$$W = KT \ln \frac{C_0}{C}, \tag{15.12}$$

где C — концентрация растворяемого вещества, C_0 — равновесная концентрация, соответствующая насыщению раствора, T — температура, K — постоянная Больцмана. Направленная энергия W связана с негэнтропией соотношением:

$$\Delta\overline{S} = K\Delta\overline{H} = \frac{W}{T}. \tag{15.13}$$

Сопоставляя (15.12) и (15.13), получаем:

$$\Delta\overline{H} = \ln \frac{C^0}{C}. \tag{15.14}$$

Таким образом, негэнтропия $\Delta\overline{H}$, существующая в системе «раствор — вещество», определяется направленностью (упорядоченностью) движения, т. е. вероятностью перехода из твердой фазы в раствор, возрастающей по мере отклонения от равновесия, т. е. с ростом $\frac{C_0}{C}$. При достижении равновесного состояния $C = C_0$ и, следовательно,

$$\ln \frac{C_0}{C} = 0 \text{ и } \Delta\overline{H} = 0.$$

Очевидно, что в таком состоянии система, с одной стороны, не обладает объективной направленностью движения, а с другой — не дает никакой информации наблюдателю, ввиду того что в состоянии равновесия в ней не происходит никаких макроскопических изменений. А пока система не достигла равновесного состояния, в ней осуществляется направленное движение, и наблюдатель может следить за происходящими в ней процессами, фиксируя изменение концентраций, давлений или температур. Этот вывод можно распространить на процессы теплообмена, химические реакции и другие направленные процессы.

Теперь предположим, что объектом наблюдений является жестко детерминированная система, но наблюдателю неизвестна заранее присущая элементам этой системы строгая взаимосвязь. Это значит, что в начальный момент наблюдений считается априорно, что все состояния равновероятны, и потому первоначальная неопределенность представлений о данной системе максимальна: $H_н\ (t_0) = H_{max}$. В результате проведенных исследований наблюдатель приходит к выводу, что при заданных исходных условиях в силу жесткой детерминации система может пребывать только в одном *(k*-м*)* состоянии, которому соответствует $p_к = 1$ (и, следовательно, $p_1 = p_2 = ...\ p_{к-1} = = p_{к+1} = \ldots = p_N = 0$). Это значит, что

$$H_н(t_к) = 0; \qquad \Delta I_н(t_к - t_0) = H_{max}; \qquad \Delta I_н(t_{к+1} - t_к) = 0.$$

Таким образом, и в этом случае дальнейшие наблюдения лишены смысла, так как, выявив присущие жестко детерминированной системе связи, наблюдатель может абстрагироваться от объекта наблюдений и предрекать все его состояния, соответствующие исходным параметрам, заданным априорно. *Новые наблюдения* над объектом понадобятся только в том случае, если в функционирование его элементов вмешалась случайность (неисправность в механизме и т. п.).

По-иному обстоит дело, если системе присущи частично детерминированные, а частично случайные связи (как это имеет место в языке). Такая система обладает неустранимой неопределенностью $H_о$ и сохраняет порядок $H_н = H_{max} — H_о$. Получив о такой системе информацию: $\Delta I_н\ (t_к — t_0) = H_н$, наблюдатель устранил неопределенность своих первоначальных представлений $H_н(t_0)$, выявил избыточную информацию и все детерминированные связи, присущие объекту. Однако и после этого он не может предрекать будущие состояния объекта в силу его стохастичных (спонтанных, мутационных) свойств. В этом случае наблюдения не прекращаются, но приобретают иной характер: теперь исследователя интересуют не интегральные статистические характеристики (распределения вероятностей и т. п.), а конкретные состояния объекта в каждый данный момент. На этом этапе наблюдатель старается выявить все спонтанные изменения, происходящие на фоне той избыточной информации, которую он может заранее предсказать. Именно так биолог наблюдает за эволюцией тех или иных популяций, социолог — за развитием общественных событий, ученый любого профиля — за соответствующей его специализации периодической печатью, каждый из нас — за ходом мыслей собеседника, лектора или автора книги, отраженным в непредсказуемом сочетании произнесенных или написанных ими слов.

С целью сопоставления рассмотренных соотношений объективной упорядоченности систем и информации наблюдений с приведенными выше высказываниями У. Эббота и У. Филдса возьмем два крайних случая:

Случай 1. Фактическая неопределенность системы максимальна: $H_{о} = H_{max}$, так как $p_1 = p_2 = \ldots = p_N$. Это значит, что в самом начале представления наблюдателя о неопределенности системы совпадают с ее фактической неопределенностью, неопределенность представлений наблюдателя отсутствует, т. е. $H_{н}(t_0) = 0$. Информация, полученная путем наблюдений, не дает в этом случае никаких новых сведений о структуре системы ($\Delta I_{н}(t) = 0$); она лишь подтверждает объективную неупорядоченность системы (наблюдатель убеждается, что $p_1 = p_2 = \ldots \ldots = p_N$).

Случай 2. Система обладает большой упорядоченностью, значения $p_1, p_2, \ldots, p_N$ резко дифференцированы. В начале исследований неизвестны значения $p_1, p_2, \ldots, p_N$; поэтому неопределенность первоначальных представлений $H_{н}(t_0)$ велика. В результате исследований наблюдатель выявил истинные значения $p_1, p_2, \ldots, p_N$, т. е. получил достаточно большое количество информации $\Delta I_{н}(t_{к} — t_0)$) позволившее ему устранить неопределенность собственных представлений $H_{н}(t_{к} — t_0)$.

Теперь сопоставим полученные нами выводы с приведенным выше высказыванием У. Филдса и У. Эббота.

Предположим, что на основании результатов многократно повторенных наблюдений построена кривая распределения вероятностей состояния элементов объекта (скоростей, координат, энергии и т. п.). Если эта кривая обладает большой дисперсией, значит система является объективно неупорядоченной и полученные о ней сведения мало уменьшили неопределенность первоначальных представлений наблюдателя. В этом случае полученная из наблюдений информация $\Delta I_{н}(t_{к} — t_0)$ будет мала (ср. случай 1). Если же в итоге исследований будет установлено, что кривая распределения вероятностей обладает малой дисперсией («острым пиком», по выражению У. Эббота и У. Филдса), то информация, полученная при наблюдении, будет достаточно велика (ср. случай 2). В самом деле, если до наблюдения исследователь не имел никаких представлений о статистических свойствах системы, то, получив информацию, по которой он строит вероятностную кривую, он может определить, какие явления наиболее характерны (им соответствует «пик» вероятностей), и утверждать, что для данной системы вероятность повторения таких-то характерных явлений велика, а отклонений от них — мала (о чем свидетельствует малая дисперсия вероятностной кривой). Таким образом, количество полученной информации $\Delta I_{н}(t_{к} — t_0)$, определяемое разностью неопределенности представлений наблюдателя до начала исследований и после их окончания, действительно оказывается обратно пропорциональным дисперсии результатов.

Рассмотренные нами соотношения объективной упорядоченности статистической системы и информации, получаемой при ее наблюдениях, полезно сопоставить с общепринятым определением коэффициента избыточности *R*. Поскольку при его определении приходится исходить из имеющихся (и далеко не всегда исчерпывающих) данных, в общем случае величина *R* зависит и от неопределенности представлений $H_н(t)$. С учетом этого обстоятельства выражение коэффициента избыточности может быть представлено в виде

$$R(t) = 1 - \frac{H_о + H_н(t)}{H_{max}}. \qquad (15.15)$$

По мере изучения статистических свойств системы

$$H_н(t) \to 0; \qquad R(t) \to R_{lim} = 1 - \frac{H_о}{H_{max}}. \qquad (15.16)$$

Соотношение (15.16) показывает, что в результате устранения неопределенности представлений $H_н(t)$ выявляется истинное значение коэффициента избыточности R_{lim}, совпадающее с общепринятым определением *R* (чтобы убедиться в этом, достаточно сопоставить выражение (15.16) с приведенным в предыдущем пункте выражением (15.4), имея в виду, что $I_ф = |H_о|$, $I_{max} = |H_{max}|$). При этом полное количество избыточной информации, содержащейся в исследуемой статистической системе ($I_{н\,max}$), определится как

$$I_{н\,max} = R_{lim}H_{max} = H_{max} - H_о. \qquad (15.17)$$

По мере устранения первоначальной неопределенности представлений об исследуемой системе обнаруживается ее избыточность, т. е. в соответствии с (15.15) уменьшается $H_н(t)$ и увеличивается коэффициент $R(t)$, который характеризует объективную упорядоченность системы (выявление которой как раз и есть цель любых наблюдений). Этот процесс иллюстрируется данными статистического анализа английских текстов, представленными в приведенной ниже таблице.

Условия расчета	Равная вероятность букв	Учет вероятностей отдельных букв	Учет вероятностей 3-х буквенных сочетаний	Учет вероятностей 8-ми буквенных сочетаний
Обозначения Шеннона	H_0	H_1	H_3	H_8
Параметры: $I_н\ (t)$	4,76 (H_{max})	4,03	3,10	1,90 (H_0)
$H_н\ (t) = I_н\ (t) - H_0$	2,86	2,13	1,20	0
$R\ (t)$	0	0,15	0,35	0,60 (R_{lim})

При ее составлении за основу приняты подсчитанные К. Шенноном значения энтропии английских текстов в алфавите из 27 букв. Значение энтропии H_s принималось за объективную стохастичность (в принятых нами обозначениях — H_0); такое допущение является вполне приемлемым ввиду малой корреляции за пределами текстового интервала длиной более восьми букв.

15.3. Информационно-системный анализ.

В настоящее время разработана основанная на идеях кибернетики общая методология исследований сложных систем. Одна из наиболее важных характерных черт системного анализа заключается в том, что он позволяет устанавливать общие свойства у совершенно различных по своей природе систем. Попытаемся показать, что объективной основой тех обобщений, которые дает системный анализ, является общность энтропийно-информационных соотношений, характеризующих структуру и состояние исследуемых систем.

Начнем с вопроса о том, каким образом соотносятся между собой понятия «информация» и «система».

Даже при беглом взгляде на эту проблему становится очевидным, что между указанными понятиями существует тесная связь. Ю. В. Сачков дает общее определение системы как «некоторой совокупности объектов (элементов), существенным образом связанных общностью (целостностью) поведения в определенных классах взаимодействий». Далее он отмечает, что простые системы могут вступать во взаимодействие, становясь элементами более сложных систем. Вполне очевидно, что в приведенном определении целостность и связность системы могут быть обеспечены только путем согласованных взаимодействий, которые обеспечиваются информационными связями между ее элементами. Поэтому необходимой компонентой любого системного анализа служит информационный подход.

Мера количества информации, определяемая нами как мера упорядоченности движения элементов системы, является одновременно и ***мерой структурно-функциональной сложности различных систем.*** Вполне очевидно, что только после того, как система достигла определенной степени упорядоченности происходящего в ней микродвижения (т. е. накопила определенное количество информации на микроскопическом уровне), в этой системе могут возникнуть сначала стабильные траектории, а затем обеспечивающая устойчивость системы как целого функциональная дифференциация отдельных ее частей. До возникновения количественной меры информации исследователи сложных систем (например, биологических) не располагали иными возможностями, кроме словесных определений, описаний и классификации их признаков. ***Мера количества информации позволяет производить точную сравнительную оценку степени сложности различных систем.***

Общность понятия «системы» позволяет объединить под этим названием такие разнородные объекты, как физические тела (системы элементарных частиц, атомов, молекул), организмы, биологические популяции, биосфера в целом, общественные структуры, языки, письменные тексты, научные теории и т. д. При этом такое объединение ни в коей мере не будет приемом формальной классификации, а, напротив, позволяет выявить целый ряд общих, присущих системам любой природы качеств и свойств. Выражением общности этих свойств служат такие понятия, как **«структура», «организация», «управление»** и т. п. При этом вполне правомерно производить сравнение структуры общественных систем со структурой нервных тканей, взаимосвязей элементов в языке со связями в живом организме и т. п. В отличие от художественных метафорических образов, строящихся по принципу аналогии отдельных черт, сопоставления различных систем оказываются многосторонними и глубокими в силу **единства природы информационных связей, составляющих фундаментальную основу всех существующих в природе систем**

Выделяя в особый класс систем **сложные управляющиеся системы**, Ю. В. Сачков относит к этому классу «системы с относительно независимым, автономным поведением подсистем (элементов) при высокой внутренней активности и избирательности, целенаправленности функционирования (поведения) системы в целом. Эти системы являются открытыми, находящимися в постоянном взаимодействии с окружением (средой) и принципиально способными решать весьма разнообразные классы задач (действовать при весьма различных обстоятельствах)». Вполне очевидно, что необходимым условием существования подобных систем будет опять-таки наличие информационных связей, но теперь уже не только внутренних, но и внешних, обеспечивающих оптимальное взаимодействие с окружающей средой.

Надо особо подчеркнуть тот факт, что использование идей кибернетики в системном анализе существенно расширило представление о возможностях обеспечения устойчивого состояния систем. **Условием устойчивости жестко детерминированных систем является неизменность («прочность») их компонент, свойств и связей. Гомеостазис обладающих сложной структурой адаптивных систем обеспечивается, напротив, лабильностью связей, существующих как между элементами рассматриваемой системы, так и между системой и окружающей средой.** Таким образом, в результате системного анализа к понятию «прочности» системы добавились понятия о ее «адаптивности», «надежности» и «лабильности». Именно этими качествами обеспечивается способность системы **самостоятельно** переходить в случае изменения **условий внешней среды** в новое **гомеостатическое** состояние. В этом и заключается принципиальное отличие «структурных» и адаптивных систем от систем жестко детерминированных, которые могут сохранять целостность и устойчивость только в условиях стабильной среды.

Из класса сложных управляющих систем следует выделить подкласс самоорганизующихся систем, способных адаптироваться к среде, изменяя не только свои параметры, но и структуру. Проблемы, связанные с исследованиями подобных систем, особенно сложны, и в этом направлении ведутся значитеьные исследования. Поскольку **структурная сложность систем определяется общим количеством сохраняемой информации и количеством уровней ее кодирования,** весьма полезно ввести в системный анализ классификацию, основанную на информационных свойствах систем. Такая классификация приводит к следующей иерархии существующих в природе систем.

I. Системы, способные сохранять информацию (т. е. направленность движения элементов) лишь на короткий период отклонения от состояния равновесия (газы, жидкости).

II. Системы, способные в течение длительного времени сохранять информацию и передавать ее родственным системам (кристаллы, магнитные или поляризующиеся среды и т. п.),

III. Системы, способные не только сохранять информацию, но и осуществлять отбор и накопление информации (живые организмы, биологические виды).

IV. Системы, способные путем целенаправленного сопоставления накопленной и вновь получаемой информации созидать новую информацию (мыслящий человек, социальные группы, системы «ЭВМ — человек»).

Как отметил А. Д. Арманд, принципиальная особенность двух последних категорий заключается в наличии специализированных функциональных органов, предназначенных для хранения, обработки и передачи информации, в то время как у двух предыдущих видов систем специализированные органы информации отсутствуют и их функции рассредоточены по всей структуре соответствующей системы.

Именно наличием ***специализированных информационных органов*** обеспечивается прогрессивное накопление информации, приводящее к развитию систем. Однако из этого вовсе не следует, что процессы развития не имеют места в системах предыдущих двух категорий. Процесс кристаллизации нельзя считать процессом развития: этот процесс детерминирован, не порождает новых качественных признаков, подобен процессу книгопечатания, в котором роль типографского набора взял на себя первый внесенный в раствор кристалл.

Однако неорганической природе присущи иные процессы: возникновение новых типов кристаллов с более сложной структурой, возникновение сложных молекул, органических соединений и т. п. Эти процессы являются звеньями единой цепи развития нашей планеты, благодаря которым в последующих стадиях возникли системы, относящиеся к более высоким категориям рассмотренной нами иерархии информационных систем.

При системных исследованиях сложных объектов основное внимание уделяется таким характеристикам, как организация, структура, взаимосвязи элементов, управление, самоорганизация. Все перечисленные характеристики можно объединить под общим названием ***функциональных характеристик***, в противовес таким ***субстанциональным характеристикам***, как масса, энергия, химический состав и т. п.

До появления кибернетики внимание всех естественных наук было сосредоточено в первую очередь на исследованиях субстанциональных характеристик. При исследованиях объектов в кибернетическом, информационном и системном аспектах субстанциональные характеристики отодвигаются на второй план. «Даже в организме,— замечает по этому поводу Л. А. Николаев,— можно заменить искусственными аппаратами и такие важные детали, как сердце, почки, легкие, сосуды, кости и т. п. ... Это убедительная иллюстрация преимущественного значения функций, слабо зависящих от материала».

Следует отметить, что предложенная К. Шенноном мера количества информации (так же как и сама категория информации), будучи одной из функциональных характеристик, имеет одно принципиальное отличие от всех остальных: ***она отражает не <u>однозначные (параметрические)</u> зависимости, а <u>многозначные (кодовые)</u> взаимосвязи элементов систем.*** Благодаря абстрагированию от субстанциональных характеристик с помощью этой меры удается прослеживать преобразование информации (перекодирование) в процессах ее накопления, хранения и передачи. Так, например, в процессе передачи телеграфного сообщения участвует множество самых разнородных по своим субстанциональным свойствам систем: текст, глаза телеграфиста, его нервные каналы, мозг, ключ, провода, приемное и передающее устройство и т. д. Анализ информационных преобразований в таком процессе может быть произведен безо всякого учета субстанциональных свойств элементов канала связи. Рассматриваются только функциональные характеристики: количество информации, способ кодирования, пропускная способность канала связи.

После открытия Второго закона термодинамики внимание исследователей было направлено преимущественно на изучение ***закрытых систем***, в которых естественно протекающие процессы неизбежно приводят к возрастанию энтропии. С появлением теории информации заметно возрос интерес к исследованиям не подчиняющихся Второму закону термодинамики и более сложных по своей природе ***открытых систем***. Поведение части этих систем подчиняется теореме Пригожина. **К их числу относятся те системы, которые, не достигая максимальной энтропии, способны неограниченное время поддерживать статистическое равновесие со средой.** Такое равновесие достигается благодаря тому, что описываемые теоремой Пригожина системы обладают **компенсирующими реакциями на воздействие внешней среды**. Другими словами, теорема Пригожина есть не что иное, как раскрытие статистического механизма принципа Ла-Шателье. Однако, несмотря на успешно преодоленную Пригожиным сложность математической формализации данного механизма, описываемая его теоремой статистическая модель еще весьма далека от модели гомеостазиса или тем более от модели эволюции структурных систем.

Если все типы систем расположить вдоль условной шкалы по степени нарастания их сложности, то системы, описываемые теоремой Пригожина, окажутся значительно ближе к термодинамически равновесным системам, нежели к сложным, структурным и многофункциональным системам, равновесное состояние которых кибернетика характеризует понятием «гомеостазис». Дело в том, что и теорема Пригожина, и статистическая теория Больцмана распространяются только на категорию гомогенных, бесструктурных систем, состоящих из множества не обладающих никакими различиями элементов. В любом случае, когда возникает различие признаков элементов («алфавит» системы) и различие их комбинаций (структуры), принципиально необходимым становится информационный подход. Именно с этих позиций анализирует физико-химическую основу биологических систем и процессов Л. А. Николаев. Полученные им на этой основе выводы полностью совпадают с оценкой роли информации, являющейся своего рода «моделью движения» в этих процессах, представленной Е.А. Седовым и приведенной нами ранее в этой работе

Как показал Л. А. Николаев, информационный подход может быть распространен на многие естественно протекающие процессы. Согласно определению этого автора, *кодом* может служить всякая *упорядоченная в пространстве или во времени последовательность воздействий.* Примером предельной упорядоченности воздействий (полной детерминации) может служить гармонический колебательный процесс.

Взаимодействие любой системы X с гармонической системой Y является частным случаем рассмотренных ранее энтропийно-информационных соотношений, описываемых условием

$$H(X; Y) < H(X) + H(Y).$$

В силу жесткой детерминации системы Y в данном случае $H(Y) = 0$, и, следовательно, $H(X, Y)<H(X)$. Другими словами, в результате гармонических воздействий система X приобрела информацию

$$\Delta I = H(X) — H(X; Y).$$

В зависимости от конкретных свойств системы X данная информация может сохраняться кратковременно (как, например, возмущения воды корабельным винтом), продолжительно (например, в песчаных дюнах) и даже «вечно» (периодические приливы и отливы морей в результате взаимного движения Земли и Луны).

Важно отметить и то обстоятельство, что все процессы информационных воздействий, начиная с простейшего случая гармонических колебаний, обеспечиваются за счет *упорядоченности* (организованности) самих *воздействующих систем* (маятника, резонансного контура, атома, поля и др.). Именно благодаря упорядоченности система приобретает способность *передачи и приема сигналов.* Так, например, упорядоченное расположение в атоме орбит электронов и их квантовых энергетических уравнений, так же как упорядоченное устройство механических маятников или колебательных электрических контуров, — основа *резонансных явлений,* которые можно рассматривать как простейшую сигнальную (информационную) связь. Пространственная упорядоченность атомов в решетке кристалла обеспечивает возможность передачи информации в процессах кристаллизации, в которых источник информации — внесенный в раствор кристалл.

Весьма интересна, на наш взгляд, данная Л. А. Николаевым информационная трактовка химических процессов и, в частности, объяснение влияния катализаторов как кодирующих систем. Оказывая влияние на ход и скорость химических процессов, катализатор, как известно, может сам не испытывать никаких параметрических изменений. Л. А. Николаев объясняет это тем, что влияние катализатора имеет чисто информативный характер: создание упорядоченности процесса либо за счет пространственной упорядоченности поверхности катализатора, либо за счет временной упорядоченности его воздействия (в цикле Кребса, в цитохромных системах и т. п.) (в данном случае следует уточнить применяемую Л. А. Николаевым терминологию: лучше **называть катализатор не кодирующей, а программирующей системой,** поскольку кодирование не приводит к увеличению количества информации, а лишь переводит заданную информацию с одного языка на другой).

Наряду с идеей широкой применимости принципа кодирования, теория информации утвердила идею *универсальности* самих кодов. Так, например, всего лишь двух признаков (знаков кода, которые можно условно обозначить как 0 и 1) оказывается достаточно, чтобы описать соответствующим — двоичным — кодом любое явление и выразить любую мысль. Чрезвычайно важно еще и то обстоятельство, что этот «язык» оказался доступным целому ряду имеющих два устойчивых состояния довольно простых устройств (реле, транзистор, магнитная ячейка и т. п.). Идея о том, что определенной последовательностью соединений этих простейших устройств достигается запоминание и обработка любой информации, в известной мере лишила покрова таинственности процесс мышления и послужила основой для построения предположений о принципах обработки информации в человеческом мозге.

Итак, информационный подход открыл новые широкие возможности анализа функциональных характеристик сложных упорядоченных систем. Однако, как это обычно случается в познавательных процессах, новый подход к явлениям обострил противоречия, выразившиеся в крайних точках зрения на сущность рассматриваемой проблемы. Не удивительно, что подобные крайности проявились и в отношении проблемы взаимосвязи функциональных и субстанциональных свойств анализируемых систем. Одна из крайних позиций в данном вопросе нашла свое отражение в высказывании У. Р. Эшби, считающего, что для успешного исследования любых систем «нужно исключить из рассмотрения два не относящихся к делу фактора. Первым из них является «материальность» — идея о том, что машина должна быть сделана из реальных материалов... Точно так же не относится к делу любая ссылка на энергию».

Против категоричности данного утверждения возражали многие ученые, в частности болгарский философ Сава Петров, потребовавший возвратить такую категорию, как субстанция (С. Петров называет ее субстратом), и высказавший предположение, что «будущее развитие науки вновь подтвердит значение субстратного анализа». Примерно такой же точки зрения придерживаются В.В. Агудов, Е.И. Свидерский, Г.А. Югай и др.

Подводя итог подобным высказываниям, Т. П. Мельников приходит к выводу, что «положение о взаимосвязи свойств структуры и субстанции объекта входит в само определение системы», ибо «субстанция содержит потенции к наиболее вероятной значимости элементов, эти значимости задают «подходящую для субстанции» структуру, а из конкретизации структуры при заданной субстанции вытекают основные внешние свойства системы как целого. Следовательно, признавая факт взаимосвязи и взаимовлияния структуры и субстанции в системах, можно рассуждать о том, какими должны быть предпочтительные структурные характеристики системы, если заданы ее внешние наблюдаемые свойства и субстанция, какова предпочтительная субстанция, если задана структура и внешние свойства, как должны быть согласованы структурные и субстантные характеристики, если заданы только внешние наблюдаемые свойства и качества системы».

На первом этапе становления категории «информация» авторов многих работ приводил в некоторое смущение явно выраженный функциональный аспект информационных процессов, ставящий категорию информации в некоторое обособленное положение относительно категорий материи и энергии и подчеркивающий ее «не-материальность». Введенная Е.А.Седовым и рассмотренная нами ранее в настоящей работе интерпретация меры количества информации, как меры упорядоченности статистических форм движения определяет место «информации» и связь этого понятия с категориями «материя», «движение» и «энергия». Становится очевидным, что сама по себе информация и не может быть «материальной», поскольку она является лишь мерой одного из присущих материи (движению) свойств. Вместе с тем очевидно, что носителями любых форм информации являются те или иные материальные объекты. Однако с методологической точки зрения обычно предпочтительно рассматривать информационные процессы в отвлечении от их субстанциональных свойств.

На первом этапе развития кибернетики некоторые ученые высказывали противоположные опасения, полагая, что новая трактовка мышления как процесса обработки сигналов, циркулирующих по цепям нейронов коры больших полушарий, лишает мысль ее «идеальности» и возрождает механицизм. Подобные взгляды высказывались, в частности, В. К. Колбановским. Данные опасения лишаются почвы в случае, если анализ процессов мышления (или других информационных процессов) учитывает противоречивость и единство их функциональной и субстанциональной сторон.

В самом деле, из-за того, что процесс мышления рассматривается как осуществление совокупностей сугубо материальных элементарных актов (возбуждений и торможений нейронов), мысль не теряет своей идеальной сущности. Образ стула, хранящийся в нашей памяти, представляет собой определенное множество материальных сигналов. И все же этот «стул» не перестает быть идеальным, поскольку в отличие от материального стула на нем нельзя сидеть.

Главное достоинство научного понятия информации заключается именно в его «емкости», позволяющей в принципе проследить функционально всю цепочку сложных явлений начиная от объективных свойств материальной системы и кончая субъективными, идеальными представлениями, обусловленными сохраняемой в чьем-то сознании информацией о той же системе. При этом **информационные процессы могут рассматриваться поочередно и в функциональном и в субстанциональном аспектах, а для полного описания их сложной природы необходимо исследовать всю совокупность их функциональных и субстанциональных свойств**.

В этой связи весьма интересно рассмотреть, каким образом тенденция к абстрагированию от субстанциональных характеристик и вычленению в исследуемых объектах чисто функциональных или чисто информационных связей изменила представление о самих «субстанциях» (можно сказать, что в данном случае в новых теоретических концепциях возникла своеобразная обратная связь).

До появления представлений о **процессе мышления как о сложном функциональном взаимодействии систем нейронов** представления о связи мышления с мозгом носили вульгарно-механистический характер: считалось, что мысли (понятия, образы) локализованы в извилинах коры. Напротив, функциональное представление о процессе мышления привело к заключению, что все, как теперь принято называть, функциональные органы, или структуры мозга, «размазаны» по коре больших полушарий (или их крупным участкам) и **объединяет их не «территориальная», а чисто информационная связь**. Можно предположить, что по мере развития интеллекта число нейронов увеличивается пропорционально увеличению площади коры больших полушарий, на поверхности которых образуется своеобразный «гофр». Этим, видимо, и объясняется выявленная ранее тенденция увеличения количества так называемых извилин по мере развития функциональных возможностей мозга. В данном случае, как и во многих других, осуществляется принцип взаимной корреляции функциональных и субстанциональных характеристик, между которыми, как правило, возникает обратная связь. Влияние функционального подхода на представления о природе «субстанции» характерно для современного этапа развития генетики. До расшифровки генетического кода идея гена связывалась с конкретными наследственными признаками: подразумевалось, что за конкретный признак ответственна сосредоточенная в хромосоме некая конкретная субстанция — «ген». Проведенные исследования дают основания утверждать, что информация о каждом признаке, связанном с теми или иными особенностями строения организма потомков (скажем, о форме крыльев или о цвете оперения птицы), может храниться не в одной, а во множестве хромосом. **Понятие «ген» стало приобретать не только субстанциональный, но и функциональный характер**. Таким образом, в этом случае, так же как при сопоставлении процесса мышления с состоянием клеток мозга, на первый план выступает ***чисто функциональная связь***.

Тот же функциональный подход преобладает в настоящее время в науке при оценках различия ***структурных уровней*** исследуемых систем. Представление о различии структурных уровней организации материи сформировалось в чисто субстанциональном аспекте. Примером этого может служить общепризнанное разграничение структурных уровней жизни: квантовый, молекулярный, клеточный, тканевый, организменный; уровни популяций, биоценоза, биосферы. В перечисленных системах градации субстанциональных и функциональных уровней полностью соответствуют друг другу, поэтому сформировавшееся при субстанциональном подходе представление об иерархии структурных уровней живой материи не претерпело никаких изменений, когда на первый план был выдвинут функциональный подход.

Однако подобное соответствие субстанций и функций отнюдь не является общим законом. Так, например, одному субстанциональному уровню может соответствовать множество функциональных уровней языка. Материальная субстанция языка в устной речи — звуковые волны, в письменной — печатные знаки. При этом одному субстанциональному уровню соответствует целая иерархия функциональных структур: связи между буквами (звуками), словами, фразами, понятиями, концепциями и т. д. Язык — средство отражения мыслей, поэтому вполне естественно, что в области мышления и построения научных теорий мы опять-таки обнаруживаем сложную функциональную иерархию, основой которой является один и тот же субстанциональный орган — мозг. Точно так же и в человеческом обществе на одном субстанциональном уровне (на уровне человека) возникает сложная взаимосвязь и иерархия функциональных социальных структур.

Только при условии четкого разграничения субстанциональных и функциональных характеристик системно-информационного анализа и их взаимосвязи можно объяснить экспериментально установленный факт независимости количества энергии, потребляемой мозгом, от интенсивности мыслительного процесса. Определено, что в период напряженной умственной деятельности (решения математической задачи, поиска художеетнеппого образа или формы и т. п.) мозг потребляет столько же энергии, сколько во время сна (около 25 ватт). Долгое время этот факт не находил объяснения именно потому, что в представлениях о процессе умственной деятельности отсутствовало четкое разграничение его функциональной и субстанциональной сторон. В результате системно-информационного подхода функционирование мозга представляется как формирование информационных связей между нейронами и циркуляция сигналов в объединяемых этими связями нейронных сетях. Эти сети взаимосвязанных нейронов и являются субстанциональной основой всех сохраняемых памятью функциональных структур (представлений, понятий, образов) независимо от того, включены ли эти функциональные структуры в мыслительный процесс, осуществляющийся в данный отрезок времени, или хранятся памятью «про запас». Отсюда становится очевидным, что участвующие в каждый момент времени в активном процессе мышления функциональные структуры составляют ничтожную долю всех сохраняемых мозгом структур; соответственно и доля энергии, расходуемой на циркуляцию сигналов в этих активизированных в данный момент структурах, настолько незначительна и по сравнению с суммарной энергией всех сохраняемых мозгом структур, что измерить изменения энергии при «включении» тех или иных структур в процесс активного мышления (например, при извлечении из памяти необходимой математической формулы) практически нельзя. Это и породило ошибочное (в силу его приблизительности) представление о независимости расхода энергии от умственного напряжения (экспериментальным подтверждением данного вывода является обнаруженное во время решения задач усиление кровоснабжения мозга, свидетельствующее об интенсификации окислительных энергетических процессов).

В этой связи следует обратить внимание на еще один весьма важный для системного анализа факт: в обладающих высокой степенью структурной организации системах сложное информационное функционирование может быть обеспечено без существенных энергетических затрат. Такая ни с чем не сравнимая по объему сохраняемой и обрабатываемой информации система, как человеческий мозг, требует для своего функционирования столько же энергии, сколько потребляет двадцатипятиваттная электрическая лампочка, дающая сравнительно слабую освещенность. Благодаря указанной особенности информационно-энергетических соотношений ничтожной энергией управляющих сигналов удается обеспечивать правильное функционирование и взаимодействие потребляющих порой огромную энергию технических систем.

Высказанные здесь общие соображения могут быть подкреплены конкретным расчетом на основе соотношения

$$\Delta S — k\Delta H = W/T,$$

где k — постоянная Больцмана, равная $1{,}38 \cdot 10^{-16}$ эрг/градус.

Переходя от термодинамических единиц к информационным (битам) путем введения множителя ln 2, представим данное выражение в виде:

$$\Delta H = W/kT \ln 2;$$

для W = 1 ватт при нормальной температуре среды ($t°$ = 20° С, $T = 273 + t°$) мы получим:

$$\Delta H = |\,\Delta I\,| = 35 \cdot 10^{12} \text{ бит.}$$

Это значит, что при КПД = 100% одного ватта энергии достаточно для питания информационной системы, содержащей $35 \cdot 10^{12}$ бит.

В равной степени справедливо и такое утверждение: получение 1 ватта направленной энергии путем превращении тепла в полезную работу соответствует увеличению информации на $35 \cdot 10^{12}$ бит.

На основании данного замечания можно поставить вопрос о введении в теорию информационно-энергетического коэффициента (правда, зависящего от температуры среды):

$$\gamma_{(t° = 20°\,C)} = 35 \cdot 10^{12} \text{ бит/ватт.}$$

Полагая, что КПД мозга приближается к 100%, мы приходим к заключению, что количество сохраняемой мозгом информации при потреблении энергии 25 ватт может достигать $\gamma = 10^{15}$ бит.

15.4. Становление вероятностных представлений (исторический обзор)

На всех этапах развития естественнонаучной и философской мысли вопрос о роли случайностей был источником разногласий между теми, кто стремился познать диалектическое развитие мира, и теми, кто такое развитие отвергал.

Сущность метафизического детерминизма, отрицающего объективную роль случайностей в происходящих в мире явлениях, наиболее четко выразил известный французский математик конца XVIII и начала XIX столетия П. С. Лаплас. До сих пор не забыта его знаменитая фраза о гипотетическом всеобъемлющем уме, который, вместив в себя все параметры, характеризующие мгновенное состояние всех частей Природы и всех одушевляющих Природу сил, смог бы выразить в математической формуле все прошлые состояния мира и предсказать по той же формуле любые события на все грядущие времена. Это высказывание Лапласа сфокусировало в себе те «фатальные» представления о мире и его познании, которые породил в XIX веке механистический детерминизм.

Последующее развитие науки опровергло этот фатальный взгляд. Стало ясно, что мир Лапласа, в котором все последующие события предначертаны формулой (пусть даже и неизвестной) — это метафизический мир. От диалектического мира он отличается тем, что в нем не может возникнуть ничего нового. Он движется, подчиняясь программе (формуле) не от простых форм к более сложным по исследуемой диалектикой восходящей спирали, а по замкнутому кругу, как заведенный единожды и навеки часовой механизм.

Диалектический мир не подчиняется этой программе. В нем происходит много явлений, которые нельзя предсказать даже с помощью фантастического числа уравнений и гипотетического всеобъемлющего ума. Ибо в диалектическом мире существенную роль играет ***случайность***.

Благодаря ***случайной изменчивости*** возникают новые формы живых организмов, и те, которые приспособлены к среде лучше, выживают в процессе отбора и передают поколениям новые признаки, которые в результате ***случайных мутаций*** приобрел сохраняемый в генах наследственный код. Благодаря *случайным ассоциациям* рождаются новые идеи, теории, взгляды. Для моделирования творческого (эвристического) мышления приходится вводить *генератор случайных данных* в детерминированную программу электронных машин. Метод этот назван «методом Монте-Карло»: заимствуя опыт игорных домов Монте-Карло, машина вдруг начинает «играть в рулетку» и с помощью случайных «очков» (стохастического сигнала) этой «рулетки» (генератора шума) находит какое-то решение даже в тех случаях, когда ей не хватает исходных данных и обычные действия в соответствии с жестко детерминированной программой не дают результата. Подчеркивая эту особенность процессов эвристического мышления, В. В. Налимов заключает: «...создавая искусственный интеллект, нужно будет ввести туда генератор случая и систему отбора, иначе мы будем иметь дело с уже совсем скучным интеллектом, построенным по принципу ветвящегося логического дерева, с четко прогнозируемым поведением».

Уже на основании этого далеко не полного перечня можно прийти к заключению, что в *диалектическом мире случайность играет принципиальную роль.*

Двадцать три века назад в несколько иной форме полемика о случайных и необходимых явлениях велась древнегреческими мыслителями. В работах античных философов уже четко наметились два подхода к оценке роли случайностей: метафизический, при котором случайность считается категорией субъективной, обусловленной неполнотой человеческих знаний; и диалектический, признающий объективную роль случайных явлений в развитии мира.

Статистическая мера количества информации и связь информации с энтропией заставляют по-новому осмыслить взаимоотношения между случайными и детерминированными явлениями и еще раз вернуться к трудам древнегреческих философов, в которых «уже имеются в зародыше, в процессе возникновения, почти все позднейшие типы мировоззрений».

Трактовка категории случайности как субъективного фактора, являющегося результатом ограниченности наших знаний, принадлежит школе древнегреческих атомистов. Такой точки зрения придерживался, в частности, Демокрит. Некоторым приближением к диалектическому пониманию соотношения порядка («гармонии», «прекраснейшего космоса») и беспорядка («мирового хаоса») является мировоззрение Гераклита, рассматривавшего развитие мира как борьбу («войну») «сходящихся» и «расходящихся» сил. Он утверждал, что, несмотря на участие «расходящихся сил», мир все же приходит к гармонии, а не ко всеобщему беспорядку («мировому хаосу»), при котором «прекраснейший космос был бы подобен беспорядочно рассыпанному сору». Этим высказыванием Гераклит отвергает абсолютизацию роли случайностей в развитии мира. Впоследствии отвергнутая Гераклитом идея превращения «прекраснейшего космоса» в подобный «беспорядочно рассыпанному сору» «мировой хаос» возродилась в теории тепловой смерти Вселенной, опиравшейся на установленный термодинамикой закон возрастания энтропии.

Дальнейшим развитием диалектического анализа проблемы соотношения необходимости и случайности являются взгляды Аристотеля, считавшего обе указанные категории объективными определениями бытия. Заслуга Аристотеля заключается в том, что он сумел преодолеть присущую многим ученым (включая и некоторых современных) тенденцию к метафизическому противопоставлению необходимости и случайности. Он считал, что, хотя эти категории и противоположны друг другу, необходимость не исключает случайности и наоборот. Согласно определению Аристотеля, к категории «случайность» относится то, что «происходит не ради какой-нибудь определенной цели, и не всегда, и не по большей части, и не в установленном порядке».

Глубокое замечание Аристотеля об отсутствии «установленного порядка» в случайных явлениях трансформировалось в современное представление о том, что **характерным и обязательным признаком стохастических процессов является <u>отсутствие алгоритма</u>, по которому можно было бы предсказывать последовательность дальнейших событий.**

Еще более поразительное совпадение с современными представлениями о диалектической взаимосвязи необходимого и случайного в материальном мире мы обнаруживаем в учении Эпикура. В противовес детерминистическому фатализму Демокрита, вытекающему из представлений атомистов о том, что **основой всех мировых событий являются механические перемещения мельчайших неделимых частиц (атомов),** Эпикур выдвигает гипотезу, предвосхитившую изучаемый современной квантовой физикой статистический характер поведения элементарных частиц. **Согласно этой гипотезе, помимо обусловленного силой тяжести «принудительного» движения атомов происходят спонтанные отклонения атомов от траектории падения.** Эпикур утверждал, что отклонения эти имеют чисто случайный характер, что происходят они «ни в определенном месте, ни в определенном времени». Вот в этих-то отклонениях, по мнению Эпикура, и кроется присущая атомам «свобода воли», нарушающая фатальную неизбежность явлений, которую отстаивал со времен Демокрита метафизический детерминизм.

Чтобы утвердить объективное существование первоисточника самодвижения материи, приводящего к непрерывному диалектическому развитию мира, Эпикуру пришлось механическому движению, обусловленному силой тяжести атомов, добавить гипотетические спонтанные отклонения, ибо *«движение падения есть движение несамостоятельности»*. И хотя с точки зрения современных квантовых представлений гипотеза Эпикура выглядит очень наивно, нельзя не признать ее философскую глубину. Ведь с ее помощью Эпикур утверждал наличие первопричины диалектического развития мира, которое пытался отвергнуть метафизический детерминизм. «В этом предположении следует особенно восхищаться тем, что оно является попыткой найти внутренний источник движения материи»,— отмечает Жорж Коньо.

Высказанную Эпикуром идею развил в своем известном произведении «О природе вещей» Лукреций Кар:

Если ж движения все непрерывную цепь образуют
И возникают одно из другого в известном порядке,
И коль не могут путем отклонения первоначала
Вызвать движений иных, разрушающих рока законы,
Чтобы причина не шла за причиною испоконъ века,
Как у созданий живых на земле не подвластная року,
Как и откуда, скажи, появилась свободная воля,
Что позволяет идти, куда каждого манит желанье,
И допускает менять направленье не в месте известном
И не в положенный срок, а согласно ума побужденью?
Ибо сомнения нет, что во всем этом каждому воля
Служит начальным толчком и по членам движенья проводит.

С. И. Вавилов также подчеркивал с позиций науки XX в. важность и глубину выдвинутых Эпикуром диалектических идей: «Нельзя умолчать о поразительном совпадении принципиального содержания идеи Эпикура — Лукреция о спонтанном отклонении с так называемым «соотношением неопределенности» современной физики... Было бы недопустимым преувеличением и грубой ошибкой считать Эпикура и Лукреция предшественниками квантовой механики, однако некоторое совпадение античной идеи с современной не является совершенно случайным. Квантование атомных состояний и энергии, несомненно, глубокими корнями связано с прерывным характером самих элементарных частиц и с диалектическими противоречиями, кроющимися в самом понятии атома».

Несмотря на «совпадение принципиального содержания идей Эпикура» с современной трактовкой соотношения случайных и детерминированных связей в явлениях окружающего мира, от Эпикура к современным научным представлениям вел отнюдь не прямой путь. Наука преодолела этот путь по спирали, причем на одном из ее участков идеи Эпикура были преданы забвению, а господствующим стал сходный с воззрениями Демокрита лапласовский фатализм. А до этого науке пришлось пережить еще и тяжесть давления теологических догматов средневековья.

Спор о соотношении случайности и детерминированности явлений приобрел в эту эпоху форму диспутов о том, предопределены ли все поступки людей предначертаниями Бога или Бог милостиво предоставляет людям некоторую «свободу воли». При этом интересно отметить, что даже в предельно «формализованном» церковными канонами религиозном мировоззрении фаталистичные детерминистские взгляды вступали в противоречие с логикой теологических концепций: за что, например, Бог наказывает грешника, если он лишен свободы воли и все его поступки, и праведные, и грешные, были заранее предопределены?

Но вот пришло Возрождение. Были открыты фундаментальные законы механики, на основе которых возродились идеи Демокрита и утвердился механистический лапласовский детерминизм. Так завершился очередной виток диалектической спирали, ибо фаталистические концепции Демокрита и Лапласа — это две «точки» восходящей спирали, расположенные на одной «вертикальной прямой».

Еще Эпикур, отвергая фатальность судьбы, вытекающую из атомистического учения Демокрита, сказал: «Лучше было бы следовать мифу о богах, чем быть рабом судьбы физиков-естествоиспытателей; миф дает намек на надежду умилостивления богов посредством почитания их, а судьба заключает в себе неумолимую необходимость». Пришедшая на смену теологическим теориям мира метафизическая концепция Лапласа тормозила развитие диалектической мысли, как в свое время религиозные догматы тормозили развитие естественнонаучных идей. Ряд ученых не только опровергали метафизический детерминизм Лапласа, но и глубоко проанализировали причины, породившие лапласовский фатализм. Одна из главных причин заключалась в том, что после длительного засилия средневековой схоластики освобожденная Возрождением научная мысль устремилась к познанию объективного мира. Для того чтобы понять свойства сложных явлений, пришлось учиться эти явления *анализировать, упрощать, расчленять.* Господствующей тенденцией научной методологии стало не исследование *взаимосвязей* явлений, а *расчленение* сложных явлений на составные части и дифференцированное изучение этих частей.

Абстрагируясь от вытекающих из этой методологии несостоятельных философских выводов, следует признать, что сама по себе данная методология на рассматриваемом историческом этапе была прогрессивной и принесла немало плодов. Так, например, путем умозрительного расчленения неразрывно связанных силы тяги и силы трения Галилей установил фундаментальный закон механики, который современникам Галилея, очевидно, казался не менее парадоксальным, чем теория относительности современникам Эйнштейна. Парадокс заключался в том, что сила лошади, запряженной в телегу, тратится вовсе не на то, чтобы придать телеге необходимую скорость, как думали до открытия Галилея.

Подчеркивая необходимость расчленения явлений как методологического приема, ряд учных отмечало: «Разложение природы на ее отдельные части, разделение различных процессов и предметов природы на определенные классы, исследование внутреннего строения органических тел по их многообразным анатомическим формам — все это было основным условием тех исполинских успехов, которые были достигнуты в области познания природы за последние шетьсот лет». Вместе с тем эта же категория ученых категорически отвергала взгляд на этот прием как на отражение объективных свойств природы, — взгляд, служивший основой метафизической трактовки окружающего мира. Метафизика абсолютизировала методологию расчленения сложных явлений и в результате упустила из виду их объективно существующие взаимосвязи. Именно отмечалось, что если метафизика права по отношению к грекам (т. е. диалектике древних*)* в подробностях, то греки правы по отношению к метафизике в целом.

Последующие исследования в области статистической физики, теории информации, биологии показали, что только наличием статистических свойств явлений можно объяснить необратимость процессов, деградацию или развитие сложных систем. Но вплоть до середины XIX века наука не искала объяснения этих явлений — она, если можно так выразиться, закрывала на них глаза. Считалось, что открытые Кеплером и Ньютоном законы небесной механики свидетельствуют о том, что все светила и планеты Вселенной вечно и неизменно обращаются по одним и тем же орбитам и нет никакой необратимости: чтобы заставить Вселенную вращаться «наоборот», достаточно во всех уравнениях перед знаком времени заменить друг на друга плюсы и минусы.

Считалось также, что, хотя организмы рождаются и умирают, живая природа в целом подчиняется одним и тем же законам, из года в год с методичностью часового механизма возобновляя свой цикл. Следуя той же тенденции расчленять природу на части, Карл Линней разложил все виды живого по полочкам, предоставив каждому виду место, закрепленное за ним на веки веков. Для развития естествознания классификация, предложенная Линнеем, явилась гигантским шагом вперед. Но базирующиеся на этой теории философские выводы явились тормозом в развитии диалектического мышления, поскольку онии утверждали неизменность животного и растительного мира.

Но вот на смену теориям, в основе которых лежит незыблемость и неизменность, приходят теории, объясняющие, как развивается мир. За неизменным движением светил и их спутников Кант усматривает возможность рождения новых космических образований и новых миров. Единство клеточной структуры растительных и животных организмов, показанное в ботанике М. Шлейденом, а в зоологии Т. Шванном, свидетельствовало о том, что полочки, предложенные Карлом Линнеем для разграничения видов, не так уж прочны. Петух, собака и клен суть не представители несовместимых форм живого, а дальние родственники, поскольку общим их прародителем является одноклеточный организм. Неизменность циклов живой природы не помешала Дарвину выявить тенденцию к эволюционному совершенствованию видов, расширив для этого рамки научных представлений за пределы жизненного опыта не только одного человека, но и всех поколений людей.

В начале XIX века мир представлялся устойчивым и неизменным, состоящим из вовеки незыблемых закостеневших форм. К концу века он во всех своих сферах представился цепью нескончаемых перемен. И вновь, как двадцать два века назад, для объяснения причин диалектического развития мира науке пришлось согласиться с тем, что в этом развитии случайность играет принципиальную роль. Правда, в отличие от времен Древней Греции наука XIX в. могла опереться на достаточно разработанный к этому времени вероятностный аппарат.

Годом рождения теории вероятностей считается 1654 год. До этого времени попытки вероятностных исчислений предпринимались Галилеем, который использовал вероятность в расчетах случайных ошибок измерений физических величин. Однако теория вероятностей как таковая начала развиваться лишь с того времени, когда удалось сформулировать и доказать несколько вероятностных теорем.

В 1654 г. кавалер де Мере предложил знаменитому Б. Паскалю найти математический метод решения одной задачи из области азартных игр. Ферма обобщил решение Паскаля, распространив его на более сложные ситуации в форме первых теоретико-вероятностных теорем.

Не следует удивляться тому, что рождению такой важной теории, как теория вероятностей, способствовал столь, казалось бы, незначительный повод, как анализ азартных игр. И по сей день азартные игры очень часто используются математикой в качестве удобных и наглядных моделей случайных процессов. Достаточно вспомнить часто встречающиеся в специальной литературе примеры, связанные с выпадением очков игральной кости, или столь важный в кибернетике метод Монте-Карло. В этой связи интересно отметить, что само слово «азарт», употребляемое теперь для выражения субъективно-психологического состояния, произошло от французского «hasard», что означает «случай».

При разработке теории эволюции Ч. Дарвин не привлекал на помощь теорию вероятностей как таковую, и все же созданная им теория насквозь пропитана духом вероятностных идей. От Дарвина и нельзя было бы требовать большего: вопрос о построении адекватных математических моделей биологической эволюции является настолько сложным, что и в настоящее время он находится лишь в начальной стадии разработки, несмотря на помощь, которую здесь оказывает аппарат теории случайных процессов и моделирование с помощью ЭВМ. Не располагая подобными средствами, Дарвин, естественно, мог лишь умозрительно представить себе механизм эволюции, в фундамент которого заложены вероятностные процессы (случайная изменчивость видов).

В этой связи нельзя не выразить восхищения по поводу другого, еще менее вооруженного фактическими знаниями, но не менее прозорливого ума. Речь идет о древнегреческом мыслителе Эмпедокле, который за двадцать три века до Дарвина тоже построил своеобразную теорию эволюции. По теории Эмпедокла эволюция осуществляется так: на одной из стадий развития мира («Сферусе») Земля породила готовые органы: ноги, сердце, глаза и т. п. Затем эти органы стали соединяться, образуя случайные комбинации, а две противоборствующие силы — Любовь и Вражда — осуществляли отбор. Все неудачные комбинации уничтожала Вражда, все удачные, гармоничные бережно охраняла Любовь. Как ни наивны подобные представления, нельзя не признать, что прообраз метода проб и ошибок здесь налицо!

Переход от общих вероятностных идей к конкретным статистическим исследованиям естественно протекающих эволюционных процессов стал возможен после создания энтропийной теории в физике и генетической теории в биологии. Обе эти теории родились в конце XIX века. В основу обеих теорий (в генетике — начиная с экспериментальных методов Г. Менделя, в физике — начиная с теоретических исследований Л. Больцмана) были положены статистические идеи. Долгое время эти теории развивались параллельно и никаких точек соприкосновения не имели. Лишь в наши дни генетика, биохимия и теория информации смогли совместными усилиями перекинуть мост от статистической меры хаоса, каковой является энтропия, к статистической мере порядка, который в форме закодированных цепочек молекул ДНК хранит в себе ген.

Созданная Л. Больцманом статистическая теория энтропии базировалась на разработанной Дж. Максвеллом кинетической теории газов, в основу которой был положен статистический анализ движения молекул. И теория Максвелла, и теория Больцмана на начальном этапе подвергались критике со стороны многих ученых, считавших их «излишними гипотезами», уводящими физиков в сторону от исследовательских задач, которые решаются на макроскопическом уровне. Об этом свидетельствуют высказывания (которые можно найти в научных статьях в журналах того времени), призывавшие к познанию «доступных явлений», а не всегда скрытой от исследователя «сущности вещей».

Однако, несмотря на критику, статистическая теория продолжала жить, причем каждый этап в ее развитии вновь подтверждал, что статистический анализ микропроцессов является главным путем познания «скрытой сущности вещей». Так, например, до статистической теории Больцмана все попытки интерпретаций физической сущности энтропии терпели полный провал. Как известно, Р. Клаузиус ввел в уравнения термодинамики функцию энтропии чисто *феноменологическим* путем, т. е. без раскрытия ее физической природы и физических свойств. Эта функция позволила установить связи между известными термодинамическими параметрами (свободной энергией, термодинамическим потенциалом, энтальпией, температурой, объемом), однако — в отличие от упомянутых параметров, имевших четкую физическую трактовку, — никакого наглядного физического толкования энтропии феноменологическая термодинамика предложить не могла. Правда, предпринимались попытки истолковать энтропию путем аналогии между работой тепловой машины и падающей воды; при этом разность температур сопоставлялась с разностью уровней наполненных водой резервуаров, аналогом количества тепла, передаваемого от источника к холодильнику, служила масса воды, а с функцией энтропии соотносились объем или вес воды. Однако более детальные исследования процессов превращения теплоты в полезную работу привели к выводу об искусственности и приближенности указанной аналогии.

Новый всплеск горячих дискуссий по поводу сущности статистических закономерностей и правомерности их привлечения для описания объективных явлений был обусловлен рождением квантовой физики и привнесенных ею в науку новых идей. Выражением крайних позиций в этой дискуссии был известный спор между Н. Бором и А. Эйнштейном (Эйнштейн утверждал, что принимать статистические трактовки в качестве объективных закономерностей равносильно признанию того факта, что все происходящее в мире определяется случайными решениями бога, принятыми в зависимости от того, сколько очков покажет подбрасываемая игральная кость).

Дискуссии о роли и методологии применения статистических методов продолжаются и в настоящее время, охватывая все более широкий круг человеческих знаний, поскольку современный этап развития науки знаменуется распространением разработанных математикой и апробированных физикой статистических методов на область кибернетических систем. Таким образом, поднятые мыслителями древности вопросы о первопричинах случайных явлений, о роли случайностей в объективных процессах, о соотношении детерминации и случайности в развитии материальных систем не решены до конца до сих пор.

15.5. Вероятностные зависимости как отражение объективных свойств материальных систем

Современная наука располагает огромным запасом данных, подтверждающих, что вероятностные взаимосвязи явлений — это одно из объективных свойств материального мира, без которого он и в самом деле превратился бы в запрограммированный, не порождающий ничего нового лапласовский мир. Привлекая на помощь теорию информации, можно использовать в качестве аргумента такой наглядный пример.

Представим себе текст, в котором полностью исключена неопределенность появления букв. В этом случае можно по части этого текста установить алгоритм, которому он подчиняется, и выписывать буквы последующих сообщений, т. е. полностью предсказывать текст. Эта ситуация может служить наглядной моделью детерминированного лапласовского мира, который не порождает ничего нового. В реальном же мире мы сталкиваемся со множеством объективно случайных, а потому и непредсказуемых однозначно явлений, для описания которых приходится использовать включающий в себя «объективные случайности» (а потому и не подчиняющийся никаким алгоритмам) язык.

Первоисточником объективной стохастичности, заключенной в самом фундаменте материи, является изучаемый квантовой физикой мир элементарных частиц. Подтверждением объективной стохастичности квантовомеханических процессов может служить принцип неопределенности, установленный В. Гейзенбергом. Этот принцип подтвержден экспериментами, показывающими вероятностный характер поведения электронов в явлениях дифракции, и многими другими экспериментальными данными.

С объективной стохастичностью явлений науке пришлось столкнуться и при исследовании генетических мутаций. Подчеркивая эту особенность генов, Дж. Хаксли писал: «Генам присуще существенное свойство жизни — копировать самих себя. Но копирование не всегда бывает точным, иногда происходит неточное самокопирование, в результате чего получаются мутации. Последние сохраняют способность самокопирования и дальнейшего мутирования. Мутации случайны или, лучше сказать, ненаправленны в том смысле, что они происходят во многих направлениях и не связаны приспособительно ни со средой, ни с общим направлением эволюции данной группы, ни с вызвавшим их агентом. Действительно, некоторые из них, по-видимому, совершенно случайны, будучи обусловлены спонтанными перестройками субатомных структур». Случайный характер возникающих в генах мутации является первопричиной стохастичного взаимодействия организмов, популяций и видов с внешней средой.

С точки зрения кибернетики отдельный живой организм или целый биологический вид есть не что иное, как сложная самоорганизующаяся система, взаимодействующая с переменными факторами внешней среды. Соответствующие заданным внешним условиям оптимальные признаки вида возникают не за счет детерминированных изменений системы, полностью адекватных внешним условиям, а благодаря статистическому механизму мутаций, аналогичному генератору случайных данных действующих по методу Монте-Карло электронных машин. В результате отбраковки множества вариантов и отбора наиболее удачных случайно возникающих признаков постепенно достигается соответствие между биологическими системами и внешней средой. Важным аспектом исследований объективной необходимости случайных связей для обеспечения жизнеспособности организмов является кибернетическая по своей сути проблема надежности биологических систем. Суть дела заключается в том, что при случайном характере внутрисистемных связей элементы системы в той или иной степени независимы, автономны, поэтому выход из строя одного или нескольких элементов не является обязательным условием для разрушения всей системы. В этом заключается принципиальное отличие вероятностных систем от систем жестко детерминированных, в которых каждый из элементов, по меткому выражению У. Р. Эшби, обладает «правом вето» по отношению ко всем остальным.

Ту же особенность вероятностных систем подчеркивал и К. Шеннон, рассматривающий в качестве наиболее надежной системы мозг. «Мозг человека или животного может служить примером очень большой и относительно надежной системы, построенной из индивидуальных компонентов (нейронов), которые надежны не только в выполнении операций, но и в тонких деталях взаимосвязи. Более того, хорошо известно, что при повреждении, несчастном случае, болезни и т. д. мозг продолжает функционировать замечательно правильно, даже если поражены его большие области. Эти факты представляют сильный контраст по сравнению с поведением и организацией вычислительных машин. Индивидуальные элементы этих машин должны быть выполнены с чрезвычайной надежностью, каждый провод должен быть соединен нужным образом, и каждая команда в программе должна быть правильной».

Теоретическому анализу преимуществ вероятностных связей с точки зрения надежности систем посвящена важная пионерская работа Дж. фон Неймана. Отмечая характерную для современной науки и техники тенденцию к использованию вероятностных связей, Ю. В. Сачков делает следующий прогноз в отношении создания высоко-надежных технических систем: «Это направление ясно выражает отказ от принципа жесткой однозначной детерминации элементов в разработке кибернетических систем. Связь между элементами строится на принципе функциональной заменяемости, функциональной избыточности: предполагается, что при выходе из строя тех или иных элементов системы выполнение их функций будет возложено на другие элементы. В идеальной перспективе биологические принципы построения кибернетических устройств ведут к принципу самораспределения функций элементов системы. Системы сами будут узнавать о возможном отказе отдельных элементов и соответственно перестраивать свою внутреннюю организацию работы, свои внутренние связи и взаимодействия... Принципиальное отсутствие однозначной детерминации элементов в подобных системах может обеспечивать непрерывность их нормального функционирования далеко за пределами той области, где действуют ныне жестко детерминированные кибернетические устройства».

Другими словами, **путь обеспечения высокой надежности — это отказ от жесткой детерминации и введение вероятностных связей в структуру проектируемых систем.**

Признание того факта, что в основу многих объективных явлений заложены случайные, вероятностные связи, далеко еще не исчерпывает всех связанных с категорией случайности научных проблем. Вслед за признанием объективной роли случайностей возникают проблемы методологического и философского плана — такие, как вопрос о предсказуемости случайных процессов, об их «удельном весе» в общей картине материального мира, а также, если можно так выразиться, об их «первоисточниках», «первоосновах». В этом плане еще не выработано общепризнанных взглядов, и это дает почву для различных трактовок, противоречащих не только друг другу, но зачастую даже самим себе.

А. Тьюрингу принадлежат слова: «Система Вселенной, как единое целое, такова, что смещение одного электрона на одну миллиардную долю сантиметра в некоторый момент времени может явиться причиной того, что через год человек будет убит обвалом в горах». Высказывание Тьюринга может быть истолковано двояко. Можно считать, что гибель человека в горах была предрешена смещением электрона. А можно исходить из того, что в цепи событий, происходивших в течение года от момента смещения электрона до обвала в горах, имело место множество случайных явлений, которые в данном случае послужили причиной обвала, зато в другом случае могли подарить горным путникам новый источник минеральной воды. Первое толкование возвращает нас к фатализму Лапласа. На самом же деле точности предсказаний помимо недостатка исходных данных препятствует объективная стохастичность явлений, для описания которых наука и разработала аппарат теории вероятностей и теории случайных процессов, выработала сложные понятия и их статистические количественные меры, назвав их «корреляцией», «информацией» и «энтропией».

С помощью статистических методов можно предречь ход самых разнообразных процессов. На основании полученных ранее статистических данных можно сделать, например, прогноз того, что в течение ближайшей недели на автомобильных дорогах Франции погибнет около двадцати человек. Таков «статистический фатализм», однако он отличается от фатализма Лапласа тем, что ни при каком количестве исходных данных не позволяет предречь, *кто же именно* в течение ближайшей недели войдет в состав обреченных статистикой двадцати человек.

Сказанное раскрывает реальное ограничение предсказуемости явлений и применимости статистических методов. Эти методы, например, не пригодны для исторической науки в том смысле, что для исследуемых ею явлений вовсе не безразлично, кто именно был в числе двадцати пострадавших и какие последствия принесла его смерть. Историю интересуют отдельные факты, привязанные к конкретному и ограниченному отрезку времени, а не статистические данные, усредненные за миллиарды лет существования Вселенной по тысячам населенных разумными жителями планет. Вместе с тем статистические методы успешно применяются современной социологией для выявления и анализа некоторых общих тенденций, характерных для тех или иных социальных групп.

Несмотря на то, что объективная необходимость вероятностных связей подтверждается всем ходом развития современной науки, и до настоящего времени в науке удерживаются взгляды, отвергающие объективный характер вероятностных связей. Некоторые считают, что вероятностные описания явлений возникают лишь потому, что в силу неполноты наших знаний мы не можем точно их описать. С этих позиций, например, анализирует некоторые высказывания Г. Паска американский ученый Г. Цопф.

Стремясь подчеркнуть основное различие между жестко детерминированными и стохастичными связями в сложной системе, Гордон Паск говорит, что в первом случае элементы системы ведут себя так, «как если бы они не могли иначе», а во втором случае — «как если бы они сами решали», как им себя повести.

Г. Цопф, пытаясь внести уточнения в эти формулировки, привносит в них элемент субъективизма и тем самым вконец запутывает вопрос. «Я уверен,— говорит Цопф,— что Паск не утверждает, что элементы должны обладать волей, хотеть и выбирать, и что они действительно таковы... Он говорит, как мне кажется, о том, что имеются некоторые невероятно сложные системы, о которых мы можем получить практически полезные знания и успешно взаимодействовать с ними, если *допустим,* что они обладают своевольной внутренней направленностью. Паск предполагает, что такой относительный подход, несмотря на то, что он не дает нам всего богатства данных и такой определенности состояния систем, которые мог бы дать нам аналитический подход, оправдан, если он дает то, что нам нужно, а не то, что мы, по нашему мнению, должны иметь».

С рассуждениями Г. Цопфа относительно условности выражения «своевольная направленность элементов системы» можно согласиться с тем уточнением, что и сам Гордон Паск вводит это понятие очень корректно («как если бы они сами решали»), т. е. не в качестве утверждения, а лишь для образной характеристики стохастических свойств. Речь тут идет, разумеется, не о «воле», а об объективной стохастичности связей элементов подобных систем. Но Цопф пытается представить суть дела так, будто все зависит только от «допущения» наблюдателя, т. е. не от объективных свойств исследуемой системы, а от того, с какой меркой наблюдатель к ней подошел.

Если же данный вопрос рассматривать с иных позиций, то можно признать, что вероятностная трактовка может быть избрана не ради удобства, а потому, что для описания объективных свойств исследуемой системы строго детерминированный («аналитический») подход просто неприменим. Тот, кто не допускает такой точки зрения, обречен выбирать между субъективным и объективным идеализмом.

Иногда обсуждают вопрос о «свободе воли» некоторых элементов материальных систем (хотя гораздо больше ученых, которые, подобно Г. Паску, употребляют это выражение лишь в качестве образной аналогии) и пытаются превратить эпикуровскую идею в некоторое знамя, утверждающее объективный идеализм. Раз частицы наделены «свободой воли», значит вероятностное описание явлений на квантовом уровне обусловлено тем, что частицы сами «решают», в какие моменты времени и в каких пределах им надлежит изменить состояние и положение в пространстве.

Отвергая подобные взгляды, Г. Цопф впадает в другую крайность. Согласно его концепции, вероятностные зависимости не есть отражение объективных свойств описываемых явлений, а только некий условный прием.

Вопреки отмеченным крайностям, в науке утверждается взгляд, что вероятность может быть не только мерой недостатка исходных данных, но и **мерой объективной стохастичности связей исследуемой материальной системы**. Четкий анализ этой тенденции провел, в частности, А. С. Кравец. «Вероятностные представления,— отмечает он,— утверждались в науке долго и мучительно. Довольно продолжительное время они не принимались всерьез в науке, считались временным костылем, которым пользуется наука за неимением лучшего. Ученые не оставляли надежд заменить в дальнейшем вероятностные законы на «истинные», как они думали, законы жесткой детерминации».

Живучесть толкования вероятностных связей как выражения неполноты наших знаний была обусловлена совокупностью исторических и психологических причин. Если в Древней Греции научные теории строились чисто умозрительно, то со времени Возрождения и вплоть до наших дней верховным судьей теоретических концепций стал **эксперимент**. **Теория** может быть признана **справедливой** только в том случае, если она способна ***предсказать***, как будет **протекать** тот или иной **реальный процесс**. Успех научных теорий в прогнозировании хода явлений создал веру в их, если можно так выразиться, «потенциальное всемогущество»: если не сейчас, то со временем наука способна все объяснить, а стало быть, и все предсказать; главной задачей любой научной теории стали считать установление однозначных зависимостей, отклонения же от них — объяснять неточностью измерений, неполнотой исходных данных и т. д. и т. п.

Отказ от сложившегося веками представления о научных теориях как о средстве однозначных предсказаний хода изученных процессов потребовал не только интеллектуальной, но и психологической ломки, пережить которую не всем ученым было легко. Как заметил Ю. В. Сачков, вероятностным представлениям, вводившим в науку принципиальную неоднозначность, «не хватало «изящества» однозначных предсказаний любых рассматриваемых связей». Подсознательное желание вернуться к «изящным» теориям, дающим возможность однозначно предсказывать ход процессов, и порождало ту самую иллюзию о неполноте и временном характере статистических теорий, из-за которой эти теории многие склонны были считать всего лишь «временным костылем».

Однако по мере развития вероятностных методов и получаемых с их помощью результатов укреплялось представление о том, что отрицание объективности вероятностных описаний явлений неизбежно приводит к выводу о существовании уравнения, о котором говорил в свое время Лаплас. Не столь существенно, знаем ли мы это уравнение сегодня, узнаем ли лишь завтра или не узнаем никогда; вместит ли исходные состояния всех частей природы один провозглашенный Лапласом гипотетически «обширный» ум или они могут быть достоянием многих (быть может, бесконечно многих) умов, — важно, что такое уравнение или система уравнений *в принципе* возможна, что в них заложена фатальная неизбежность «программы», которой подчиняется наш мир.

В действительности подобной «программы» не существует. «В ходе развития действительности,— пишет Ю. В. Сачков,— происходит не только выбор определенной возможности из множества заданных, но и возникают существенно новые возможности, а следовательно, не все возможности определены ранее действительностью. Новое всегда в чем-то не определено... Реальная жизнь всегда вносит коррективы, иначе следует признать абсолютную предопределенность настоящего прошлым с неизбежным сползанием в фатализм». Другими словами, наука стоит перед дилеммой: либо признание феномена объективной стохастичности явлений, либо возврат к фатализму Лапласа. Третьего не дано.

Польский ученый Мариан Смолуховский впервые поставил эксперимент, говорящий в пользу объективной природы случайных явлений; ему принадлежит и заслуга создания **теории флюктуации**. Эксперимент, придуманный Смолуховским, отличался остроумием и изяществом. До этого эксперимента считалось, что голубой цвет неба обусловлен наличием в воздухе посторонних частиц. Чтобы опровергнуть это объяснение, Смолуховский показал, что заключенный в специальной трубе идеально очищенный воздух тоже кажется голубым. Причиной этого являются стохастичные флюктуации плотности воздуха, обусловленные тем хаотическим движением молекул, которые описал Джемс Максвелл.

Оказалось, что доказательство объективности случайных явлений в буквальном смысле «витало в воздухе». «Люди всегда имели живое доказательство перед глазами, нужно было только научиться читать в книге природы», — писал Смолуховский. «Поскольку дело касается применения в теоретической физике, все теории вероятностей, которые рассматривают *случайность как непознанную частичную причину,* должны быть заранее *признаны неудовлетворительными.* **Физическая вероятность события может зависеть *только от условий, влияющих на его появление,* но не *от степени нашего знания».***

Исходя из представления об объективном характере вероятностных явлений и их связи с внешними условиями, В. А. Фок связал понятие «вероятность» с философской категорией «возможность», определив вероятность как потенциальную возможность поведения объекта. «Вероятность того или иного поведения объекта в данных внешних условиях,— пишет Фок,— определяется внутренними свойствами данного индивидуального объекта и этими внешними условиями; это есть численная оценка потенциальных возможностей того или иного поведения объекта. Проявляется же эта вероятность в **относительном числе** осуществившихся случаев данного поведения объекта; **это число и является ее мерой.** Таким образом **вероятность относится к отдельному объекту и характеризует его <u>потенциальные возможности</u>**...».

Аналогичную трактовку вероятностного описания траекторий элементарных частиц как некой разновидности «потенций» дает В. Гейзенберг. Подчеркивая объективный характер вероятностных связей в таких физических явлениях, как флюктуации, броуновское движение, изменение энтропии во времени и т. п., Н. С. Крылов отмечает, что «вероятностный характер полученных результатов измерений является абсолютно *достоверным фактом опыта,* не менее достоверным, чем вероятностный характер рядов испытаний, полученных при любом другом, сколь угодно обоснованном приложении теории вероятности... Полученные при таких измерениях ряды результатов измерений обладают, следовательно, **общим свойством всех вероятностных рядов — отсутствием алгоритма**, который определял бы результаты последовательных измерений. Сколько угодно сложная формула принципиально не может описать последовательный ход результатов измерений величин, подчиняющихся вероятностному закону распределения свойств».

Отсутствие указанного алгоритма, позволяющего предсказать последующий ход событий, как раз и является выражением объективной стохастичности случайных процессов. Именно неопределенностью, непредсказуемостью они отличаются от процессов, обусловленных жестко детерминированными связями, благодаря которым «при заданных внешних воздействиях начальное состояние однозначно определяет все дальнейшее движение системы». Объективный характер иррегулярности единичных явлений в стохастических процессах, выражающийся в отсутствии алгоритма, подчеркивает также А. С. Кравец: «**Невозможность алгоритмического описания вероятностной системы** нельзя представлять себе как нашу объективную беспомощность, как ограниченность наших познавательных возможностей. Отсутствие детерминированных алгоритмов для вычисления отдельных результатов испытаний над системой представляет некоторый объективный закон вероятностной системы. Невозможность нахождения того, чего в действительности нет, нельзя ни в коей мере считать проявлением нашего незнания или слабости наших методов познания».

Развивая ту же мысль об объективности вероятностных связей в целом ряде процессов, Б. Рассел приходит к выводу, что вопреки представлениям Лапласа в этих процессах «незнание все-таки не включено в понятие вероятности, которое имело бы тот же смысл для всеведущего существа, как и для нас». Больше того: «отсутствие всяких неизменных правил в отношении элементов вероятностной системы переходит в свою противоположность и *само становится некоторым правилом,* обязательным для вероятностных систем». *Вот почему для установления объективной стохастичности тех или иных процессов часто приходится специально доказывать с помощью* *многократно повторяемых выборок, что в них отсутствуют регулярная повторяемость и какая-либо жесткая детерминированная связь. Только в этом случае к исследованию этих явлений становится применимым статистический подход.*

Все сказанное выше относительно объективного проявления вероятностных связей в стохастичных процессах вовсе не исключает случаев, когда введение вероятностей в те или иные расчеты и в самом деле обусловлено недостатком исходных данных. Так, например, с помощью определяемого статистическими методами так называемого среднего роста («математического ожидания» роста) можно довольно точно определить суммарную потребность в материале для спецодежды группы рабочих, если известно количество рабочих, но неизвестны размеры одежды каждого из них. В этом случае расчет содержит определенную долю субъективизма, обусловленную неполнотой предварительных сведений: отсутствием списка размеров одежды конкретных рабочих, неточностью выявления статистического «среднего роста».

Задача статистического исследования тех или иных конкретных явлений обычно как раз и заключается в том, чтобы из задаваемых априорно распределений вероятностей исключить субъективный элемент. Так, например, при первых попытках оценки количества информации, содержащейся в письменных текстах, априорно предполагалась равная вероятность появления всех букв (мы уже отмечали, что именно так предлагал определять количество информации предшественник К. Шеннона Р. Хартли, используя формулу $I = \log N$, где N — число букв алфавита). Однако последующие статистические исследования письменного языка позволили не только обнаружить различия вероятностей появления разных букв в текстах, но и учесть статистически взаимосвязь (корреляцию) между буквами и словами. В результате исключения субъективного фактора (неполноты исходных сведений) при оценке статистической неопределенности русского текста с помощью функции

$$\sum_i p_i \log p_i$$

было установлено, что неопределенность эта меньше, чем заданная априорно (5 бит на букву), приблизительно в 5 раз.

15.6. Диалектическая природа вероятностного анализа

Решение вопроса о том, в какой мере заданные значения вероятностей или неопределенности (энтропии) отражают объективные свойства процессов, а в какой мере являются отражением неполноты априорных знаний, сопряжено с серьезными трудностями, отражающими диалектический характер взаимоотношения реальности и ее познания в науке.

Методология анализа стохастичных процессов весьма выпукло отражает указанные трудности. Подчеркивая различия между диалектическим и метафизическим способами мышления, ряд ученых говорили о том, что при метафизическом подходе «какая-нибудь вещь, какое-нибудь отношение, какой-нибудь процесс либо случайны, либо необходимы, но не могут быть и тем и другим». Многие ученые до сих пор пытаются по принципу «или — или» противопоставить случайные явления необходимым. Можно указать три категории таких противопоставлений:

1. Объективное существование признается только за необходимыми связями, а случайные связи относятся к категории субъективных факторов.
2. Абсолютизируются случайные связи и отрицается объективное существование связей необходимых.
3. Признается объективность и случайных и необходимых связей, но отрицается возможность их совмещения во времени и в пространстве.

Примером отрицания объективности случайных связей может служить рассмотренная в предыдущем пункте позиция Г. Цопфа. Как уже отмечалось, указанная позиция бытует в науке еще со времен Демокрита, а дальнейшее ее развитие привело к концепции механистического лапласовского детерминизма.

Не нова и тенденция абсолютизации роли случайных связей, хотя возникла она значительно позже в связи с развитием вероятностных методов и расширением круга решаемых этими методами задач. Чтобы проверить на практике действенность введенного в математику Паскалем и Ферма понятия вероятности, английский математик К. Пирсон подбрасывал монету 24000 раз. Убедившись, что закон равновероятного выпадания обеих сторон монеты выполняется тем точнее, чем большее число раз повторяется опыт, Карл Пирсон пришел к заключению, что «необходимость принадлежит лишь к миру понятий».

Так возникла новая метафизическая тенденция, полюсная по отношению к традиционному отрицанию объективности случайных явлений. Дальнейшее развитие этой метафизической тенденции привело к теории тепловой смерти Вселенной, утверждающей, что безраздельное господство случайностей — это будущее, ожидающее весь мир.

Убедительным опровержением концепций, противопоставляющих необходимость случайности, может служить устойчивость некоторых характеристик случайных процессов. Что означает, что вероятность некоторого события равна, например, 0,1? Это означает, что данное событие повторится примерно 10 раз из 100 выборок, или 100 раз из 1000, или 1000 раз из 10 000. При этом согласно установленному теорией вероятностей закону больших чисел соотношение 10/100, или 100/1000, или 1000/10 000 будет подтверждаться опытом тем точнее, чем больше произведено проб. Не есть ли это вероятностная форма проявления необходимости?!

С точки зрения диалектики упомянутый опыт К. Пирсона не опровергает необходимости. Многократно подбрасываемая монета подтверждает существование необходимости, выраженной в вероятностной форме и заключающейся в том, что вероятностные соотношения выполняются тем точнее, чем больше произведено бросков. Стоит лишь допустить, что такой необходимости объективно не существует, как понятие вероятности теряет какой бы то ни было смысл. В самом деле, можно ли установить значение вероятности события, если в первой тысяче выборок оно повторилось 3 раза, во второй тысяче — 150 раз, а в третьей — 995 раз? Закон больших чисел свидетельствует о том, что лежащее в фундаменте теории вероятностей и теории случайных процессов основное понятие (вероятность) уже совмещает в себе два диалектически противоположных начала: с одной стороны, оно говорит о том, что характеризуемое им явление *случайно,* а с другой — утверждает *необходимость повторения этого случайного явления* определенное наперед заданное число раз. Так, сквозь «толпу» из N случайных событий пробивает себе дорогу необходимость, заключающаяся в том, что каждое i-е событие неизбежно должно повториться $p_iN=n_i$ раз (полагая, опять-таки, что $i = 1, 2, 3 \ldots, \sum_i p_i = 1$).

При этом чем больше «толпа случайностей» (число N), тем четче проявляет себя необходимость (точнее выполняется соотношение $n_i = p_iN$).

В этом и заключается диалектика: случайность и необходимость, будучи по сути своей противоположны, в реальных явлениях неразрывно связаны между собой. Этой связи не улавливали те ученые, которые воспринимали развитие вероятностных методов в науке как доказательство безраздельного господства случайностей, исключающих всякую необходимость.

В несколько видоизмененной форме тенденция к крайностям в оценках роли случайных и детерминированных явлений существует и в науке наших дней. Так, например, анализируя процессы эволюции или творческого мышления с кибернетических позиций, многие авторы склонны абсолютизировать роль случайностей в процессах поиска новых решений. К подробному анализу «неометафизической» тенденции и вытекающих из нее выводов мы вернемся ниже.

Опыт исследований стохастичных процессов, приобретенный современной наукой, со всей очевидностью демонстрирует плодотворность диалектического анализа единства случайных и детерминированных связей явлений. Как отмечает М. Лоэв, в течение почти двух столетий понятие вероятностной связи трактовалось как синоним независимости явлений. В начале нашего веке благодаря исследованиям А. А. Маркова в теорию вероятностей прочно внедрились представления о *частичной* зависимости, получившей название *корреляции* случайных величин. Оказалось, что в ходе случайного процесса мгновенное значение изменяющегося по случайному закону параметра не полностью произвольно — оно в той или иной степени зависит либо от других случайных параметров (взаимная корреляция), либо от того, каким был этот параметр несколько мгновений назад (автокорреляция). По величине введенных теорией случайных процессов коэффициентов корреляции можно судить, с какой степенью жесткости осуществляется в том или ином случайном процессе частично детерминированная связь.

В свете идей кибернетики наличие автокорреляции трактуется как одна из форм проявления памяти стохастичных систем. В самом деле, если значение того или иного изменяющегося по случайным законам параметра зависит от того, каким был тот же параметр несколько мгновений назад, значит, система, в которой этот процесс протекает, обладает кратковременной памятью, регулирующей случайный процесс. Это не что иное, как еще одна из форм проявления взаимосвязи между сохраняемой (запоминаемой) информацией и упорядоченностью движения материальных систем.

Благодаря исследованиям Маркова вместе с понятием корреляции возникло понятие *стационарности случайных процессов,* выражающееся в неизменности усредненных статистических характеристик. В ходе стохастичных процессов те или иные параметры приобретают различные случайные значения в каждый данный момент; вместе с тем усреднение этих параметров за длительный интервал времени обнаруживает, что и при случайных изменениях соблюдается свой порядок. Выражением этого порядка является неизменность во времени таких характеристик, как распределение вероятностей, коэффициент корреляции, математическое ожидание, моменты распределения различных порядков и т. п. Развитие теории случайных процессов явственно показывает, что понятие «чистой случайности» относится к области теоретических абстракций, поскольку все реальные случайные процессы в той или иной степени детерминированы.

Изучение соотношения случайности и необходимости, их диалектической взаимосвязи очень важно для науки. Так, например, успехи квантовой физики в немалой степени связаны с тем, что ее теоретические концепции базируются на учете взаимосвязей этого рода. В квантовой механике вероятностный характер поведения микрочастиц, описываемый волновым уравнением Шредингера, прекрасно уживается с детерминированными характеристиками тех же микрочастиц: квантовыми числами, спинами, странностями и т. п. Именно эти параметры характеризуют упорядоченность квантовомеханических процессов, регламентируют стохастическое движение микрочастиц.

Другим примером неразрывного сочетания случайных и детерминированных связей могут служить модели так называемых эмергентных (творческих, создательных) процессов, создаваемые с помощью электронно-вычислительных машин. Практика создания программ для ЭВМ, имитирующих элементы эвристического (творческого) мышления, со всей очевидностью показала, что для решения задач эвристики в жестко детерминированные процедуры должен быть обязательно включен элемент случайности (стохастичный сигнал). В этом и заключается смысл метода Монте-Карло, а также метода «проб и ошибок». При случайном поиске оптимальных решений с помощью электронной машины каждый удачный шаг неизбежно сопровождается множеством ошибочных проб. Сразу напрашивается аналогия между процессом, протекающим в действующей по методу «проб и ошибок» электронной машине, и биологической эволюцией, осуществляющейся посредством естественного отбора, т. е. опять-таки ценой множества неудачных попыток.

Стоит лишь сопоставить эволюцию с машиной, работающей по принципу «проб и ошибок», чтобы обнаружить, что в обоих случаях основой является взаимодействие трех компонентов: (1) случайных данных (стохастические сигналы, генерируемые в электронной машине, мутации генов в живых организмах); (2) ограничения пределов случайного поиска (коррективы поиска и запоминание оптимальных решений в машине; сохраняемый поколениями генетический код) и (3) отбора рациональных случайных данных (в машине—по критерию приближения результатов решений к условиям оптимума, в природе — по адекватности новых признаков условиям внешней среды).

Согласно научной генетике ограничения случайной изменчивости организмов, обеспечивающие сохранение поколениями живых организмов главных признаков вида, обусловлены независимостью программы, заложенной в генетическом коде, от сигналов, поступающих в организм из внешней среды. Внешними факторами, приводящими к мутациям генов, являются не информационные, а энергетические воздействия (например, жесткие излучения), дающие не адекватные изменения, а случайные замены нуклеотидов в молекуле ДНК. В результате в самом организме или его потомстве возникают новые свойства — признаки, рассматриваемые в дарвиновской теории как результат случайной изменчивости. Затем вступает в действие механизм естественного отбора, с помощью которого природа «решает», какой из случайно возникших признаков в данных условиях оптимален и достоин быть закрепленным в потомстве.

Аналогия между биологической эволюцией и поиском эвристических решений, осуществляемым с помощью ЭВМ, имеет глубокие корни. Она обусловлена тем, что оба рассматриваемых процесса относятся к процессам диалектического развития, приводящим к возникновению новых признаков, качеств и свойств. В основу обоих процессов заложено диалектическое единство двух противоположных начал: случайного поиска и детерминированного отбора.

Данный принцип, по всей видимости, является универсальным для всех видов эмерджентных процессов, будь то творческий поиск решения научной проблемы или творчество самой природы, воплощающееся в эволюционном процессе. Вместе с тем наряду с общностью принципов эвристических программ ЭВМ и естественно протекающих эмерджентных процессов между ними существует и немало различий. Так, например, сопоставление результатов исследований свойств клетки коры головного мозга (нейрона) с логическими элементами, используемыми в схемах ЭВМ, обнаруживает существенные различия между принципами обработки информации в головном мозге и в машине. Электронный логический элемент имеет два устойчивых состояния, соответствующих двум знакам двоичного кода (0 или 1) или двум логическим возможностям («да» или «нет»). Сигналы же, вырабатываемые нейроном, нельзя свести к двум указанным знакам, так как существенную роль в нервных процессах играют их флуктуации по интенсивности и частоте. Этот экспериментально установленный факт послужил причиной возникновения гипотезы, согласно которой в основе обработки информации в коре головного мозга лежит не столько двузначная, сколько вероятностная («серая») логика.

Если двузначная логика на каждом этапе дает одно из двух заключений («да» — «нет») с вероятностью 1, то вероятностная логика допускает любые значения вероятностей указанных утверждений. В результате возникает логика, в которой утверждения распределены по шкале вероятностей, т. е. охватывают всю гамму оттенков «серого цвета», распределенных между двумя полюсами:

«белым» ($p_{«да»}=1$, $p_{«нет»}=0$) и «черным» ($p_{«нет»}=1$, $p_{«да»}=0$). При этом на смену «жестким» оценкам исследуемых явлений по принципу «или — или» приходят, оценки, допускающие в качестве промежуточного или конечного утверждения заключения типа «да, частично» («да» — с вероятностью *p*); «и да и нет» («да» — с вероятностью *p*, «нет» — с вероятностью 1 — *p*).

Как уже отмечалось, факт реализации «серой» логики в мозге пока является гипотетическим. Правомерность существования вероятностной логики наряду с двузначной логикой, основанной Аристотелем и математизированной Булем, обоснована в ходе дальнейшего развития электронно-вычислительной техники и связанных с ней теоретических направлений. Можно утверждать, что элементы подобной логики содержатся во всяком вероятностном описании, ибо установление того факта, что некое свойство объекта может реализоваться с вероятностью *p* или не проявиться с вероятностью (1 — *p*), означает признание за данным объектом права быть «и тем и другим».

Метафизические выводы «или — или» устанавливают «hard and fast lines» (абсолютно резкие разграничительные линии) явлений. Возникающие при вероятностных оценках заключения типа «и да, и нет» переводят «друг в друга неподвижные метафизические различия», позволяют исследовать процессы диалектического развития, в которых «все различия сливаются в промежуточных ступенях, все противоположности переходят друг в друга через посредство промежуточных членов». Вот почему именно вероятностные методы позволяют современной науке исследовать единство противоположностей волновых и корпускулярных свойств микрочастиц, процессы развития биологических видов, мышления, языка и прочих самоорганизующихся систем.

В самом деле, до тех пор, пока вопрос о закономерностях развития биологических видов наука пыталась рассматривать с позиций «жесткой» логики, научная мысль бесплодно вращалась в порочном кругу. До появления теории Дарвина периодически возникали споры между представителями двух метафизических школ. Одна из них утверждала, что в процессе развития организм не испытывает никаких качественных изменений, потому что в оплодотворенном яйце уже существует в миниатюре весь будущий организм; эта точка зрения нашла свое выражение в теории *преформизма.* Другая метафизическая концепция — теория *эпигенеза;* согласно этой концепции признаки, возникающие при развитии организма, полностью соответствуют условиям среды. Отсюда следует, что каждое поколение заново «вырабатывает» все полезные для него признаки, а потому отсутствуют наследственность и накопление полезных признаков.

Взглянем на вопрос с позиции «вероятностной логики». Существует ли изменчивость организмов? Теория эпигенеза дает ответ «да», теория преформизма дает ответ «нет». Оба ответа категоричны и однозначны, поэтому взаимно исключают друг друга. И в этом ошибочность обеих теорий. Теория эволюции дает ответ типа «да, частично». Изменчивость есть, но не безграничная, а в определенных рамках (в пределах заложенной в генах «программы»), сохраняющих главные признаки биологических видов и в то же время допускающих возможность их эволюции. И в этом заключается диалектическая природа эволюции, рассматриваемой со времен Дарвина как результат борьбы двух противоположных начал: стремления к изменчивости и стремления к сохранению основных признаков вида, которые несет в себе наследственный генетический код.

Аналогичный этап перехода от метафизических детерминистских трактовок явлений по принципу «или — или» к признанию вероятностной природы исследуемых объектов пришлось пережить физике элементарных частиц. История возникновения квантовой физики — еще один пример перехода от «двузначной» к «вероятностной» логике, характерного для современной науки. Анализируя фундаментальные выводы квантовой физики, легко обнаружить связь между присущими вероятностному анализу заключениями типа «и да, и нет», «да, частично», с одной стороны, и результатами исследований элементарных частиц, с другой.

Находится ли частица в данной точке пространства? На этот вопрос не может быть дан присущий «детерминистской» логике однозначный ответ. Квантовая механика дает ответы «да, частично», «и да, и нет»: вероятность того, что частица находится в данной точке пространства пропорциональная квадрату амплитуды шредингеровской ψ-волны.

С подобными ситуациями при исследовании процессов на квантовом уровне физика сталкивается буквально на каждом шагу. Так, например, световой луч, проходящий сквозь кристаллическую решетку, поляризуется частично в направлении, параллельном оптической оси кристалла, а частично — перпендикулярно этой оси. Эту особенность квантовой физики отметил один из ее создателей П. Дирак: «В классическом смысле слова нельзя представить себе, что система находится частично в одном состоянии, а частично в другом и что это эквивалентно тому, что система целиком находится в некотором третьем состоянии. Здесь вводится совершенно новая идея, к которой надо привыкнуть».

Развитие современной науки, и в первую очередь квантовой физики и теории самоорганизующихся систем, со всей убедительностью показало, что заключенная в «вероятностной логике» субъективная диалектика является отражением объективной диалектики взаимодействия случайных и детерминированных связей, лежащего в основе всех процессов развития материальных систем.

Ретроспективный взгляд на развитие научного мировоззрения с позиций современной науки позволяет сделать следующий вывод: ***обобщением анализа, основанного на детерминистических трактовках, являются законы сохранения*** (например, установленные классической физикой законы сохранения энергии, массы, импульса, закон движения небесных тел по неизменным орбитам и т. п.). ***Статистические закономерности, для исследования которых наука неизбежно обращается к статистическим мерам (корреляции, информации, энтропии) — фундаментальный источник всех изменений, происходящих в реальных процессах*** (будь то превращение энергии из одной формы в другую, взаимные превращения элементарных частиц, развитие самоорганизующихся систем, образование из космической пыли новых космических тел и т. п.).

В заключение следует отметить, что в настоящее время для описания различных сущностей, их анализа и синтеза используется нечеткая математика, в том числе, нечеткая логика, о чем будет идти речь в последующих книгах настоящей работы.

16. Диалектика случайных и детерминированных связей в эволюционных процессах

16. 1. Уточнение понятия развития в терминах теории информации

Рассмотренные ранее примеры показывают, что процессы развития обусловлены увеличением количества информации, накапливаемой и сохраняемой в структуре развивающихся систем. Данная точка зрения, получила дальнейшее развитие. Вместе с тем из результатов работ ученых следует, что внедрение методов теории информации в исследования процессов развития сопряжено с целым рядом методологических трудностей, к числу которых следует отнести отсутствие определения самого понятия «развитие» и четкого разграничения его форм. В данном разделе делается попытка, во-первых, дать определение понятия развития в терминах теории информации и теории случайных процессов, а во-вторых, выявить и сформулировать в тех же терминах некоторые общие свойства этих процессов.

Среди направлений современной науки биология представляет ту область, для которой вопросы методологии исследований процессов развития особенно актуальны. Именно поэтому в биологии раньше, чем в других областях, назрел вопрос о необходимости уточнения идеи развития и разграничения видов развития, поскольку до настоящего времени одним и тем же термином «развитие» обозначался нередко и филогенез, и онтогенез. Между тем имеется принципиальная разница между филогенезом и эмбриональной стадией онтогенеза и морфогенеза. Эмбриональное развитие зародыша из зиготы (онтогенез), в частности развитие и дифференциация форм его органов (морфогенез), осуществляется в соответствии с заложенной в ДНК детерминированной программой (согласно современным научным воззрениям вероятностная трактовка — одна из форм проявления детерминизма. В связи с этим необходимо внести уточнение: в дальнейшем изложении термины «детерминированные связи», «детерминированная программа» и т. п. мы будем употреблять как антонимы терминов «вероятностные связи», «вероятностные программы» и т. д., относя прилагательное «детерминированный» («детерминированные») к жестким связям, при которых начальное состояние и заданный комплекс внешних условий однозначно предопределяют последующие состояния элементов систем).

В этой тадии находящийся в оплодотворенной яйцеклетке зародыш может быть отнесен к категории если не полностью, то во всяком случае преимущественно *закрытых* для внешней информации детерминированных систем (птенец в яйце). Этим эмбриональные процессы развития коренным образом отличаются от развития популяций и биологических видов (филогенеза), поскольку популяция (или вид) является открытой информационной системой, развитие которой определяется не только изначальной программой (генофондом), но и взаимодействием случайной изменчивости и естественного отбора, обусловленного воздействием внешней среды.

В ходе развития популяций и видов происходит образование новых полезных признаков, не предусмотренных изначальной программой. Благодаря этому эволюция видов представляет собой один из наиболее ярких примеров движения по восходящей линии от простых форм к более совершенным и сложным, и потому именно эволюция (а не детерминированные процессы эмбрионального онтогенеза и морфогенеза) соответствует тому определению развития, которое вкладывает в это понятие диалектический материализм.

Как было показано ранее, всякое развитие может характеризоваться увеличением количества информации, сохраняемой развивающейся системой. При этом статистическая мера количества информации служит мерой объективной упорядоченности движения элементов системы, а содержательная сторона информации заключается в сохранении и передаче «моделей» (алгоритмов) движения элементов систем. Однако условие накопления информации для развития необходимо, но недостаточно. В этом сможет нас убедить следующий пример.

В книге М. Аптера «Кибернетика и развитие» рассмотрены некоторые процессы развития, моделируемые на электронных машинах с использованием детерминированных программ. На первый взгляд кажется, что эти модели опровергают высказанную ранее точку зрения, согласно которой основой всяких процессов развития должен являться не детерминированный, а вероятностный (стохастический) механизм. На самом же деле кажущиеся противоречия между примерами Аптера и точкой зрения, изложенной ранее, носят чисто терминологический характер. Дело в том, что термином «биологическое развитие» М. Аптер обозначает процессы роста и дифференцировки, причем, следуя П. Вейсу, предлагает обозначать термином «рост» увеличение числа элементов одного сорта, а термином «дифференцировка» — увеличение числа сортов. Именно такие процессы (рост, рост с дифференцировкой) М. Аптер моделирует с помощью детерминированных программ на ЭВМ. Но коль скоро речь заходит об *эволюционном развитии* (М. Аптер именует этот процесс просто *эволюцией* с целью подчеркнуть отличие ее *от биологического развития),* то тут и сам Аптер признает необходимым ввести в программу электронной машины ту или иную стохастическую модель (в частности, химическую машину Г. Паска, в которой в качестве стохастической модели используются нити случайной конфигурации, возникающие вследствие электролиза и выделения ионов металла, растворенного в кислоте). Очевидно, что в принятой нами терминологии только понятие эволюционного развития соответствует диалектическому представлению о развитии, в то время как детерминированные процессы роста и дифференцировки по заложенной в ЭВМ или в ген программе не дают ни новых признаков, ни новых связей. Они являются лишь повторением информации, заложенной первоначально в программе, и в этом отношении их можно сравнить с процессом печатания книг, не порождающим ни новых теорий, ни новых идей.

Следует подчеркнуть, что рассматриваемые М. Аптером детерминированные модели имитируют развитие лишь *информационно закрытых систем.* Это не что иное, как лапласовский мир в миниатюре, мир, в котором не происходит ничего, что не было бы предусмотрено изначальной программой, определяющей весь ход происходящих в этом мире событий.

Между тем основой развития реального мира являются не закрытые, а открытые для информации системы, поскольку механизм естественного отбора основан на сопоставлении признаков, возникающих в результате внутренней случайной изменчивости, с информацией, поступающей из внешней среды. Согласно существующим воззрениям аналогичный принцип отбора распространяется не только на биологические системы, но и на объекты неорганической природы. В результате возникают устойчивые к внешним воздействиям системы, адаптация которых обусловлена их свойствами, удовлетворяющими обобщенному принципу Ле Шателье. Из этого следует, что одним из условий диалектического развития является наличие открытых информационных каналов, по которым осуществляется связь системы с внешней средой.

Сопоставление детерминированных моделей М. Аптера с процессами эволюции приводит к мысли о том, что не всякое накопление информации тождественно диалектическому развитию. К категории диалектического развития относятся только такие процессы, в которых в результате взаимодействия происходит не многократное повторение и размножение известных заранее признаков, а появление новых. Так, например, в процессе развития в результате взаимодействия нескольких сложных систем может возникнуть новая суперсистема, в которую каждая из исходных входит как составная часть. Новая суперсистема будет «системой», а не «механической смесью» только в том случае, если дело не ограничивается простым «совмещением» вошедших в нее исходных систем. **Между исходными системами должны возникнуть новые связи, в результате чего количество информации, сохраняемой вновь возникшей суперсистемой, будет превышать суммарное количество информации исходных систем.** С учетом сказанного ***развитием***, в диалектическом понимании этого термина, следует считать лишь такие процессы, в которых ***увеличение количества информации опережает как увеличение массы системы (М), так и рост числа составляющих ее однородных элементов (N).*** На языке математики эти условия выражаются соответственно как:

$$\frac{\partial^2 I}{\partial M^2} > 0, \qquad \frac{\partial^2 I}{\partial N^2} > 0.$$

Анализируя роль естественного отбора в процессах образования систем, А. Л. Тахтаджян делит отбор на две категории: *(1)* ***матричный отбор****, приводящий к многократному образованию идентичных систем; и (2)* ***эмерджентный отбор****, приводящий к образованию новых связей и соответствующих новых качественных признаков систем*. Тахтаджян считает, что в отличие от эмерджентного (творческого, созидательного) отбора количество информации при матричном отборе не увеличивается. По нашему мнению, здесь следует внести уточнение: количество информации возрастает как при матричном, так и при эмерджентном отборе, только в первом случае количество информации возрастает линейно $\left(\frac{\partial I}{\partial N} > 0; \frac{\partial^2 I}{\partial N^2} = 0\right)$, а во втором — прогрессивно $\left(\frac{\partial^2 I}{\partial N^2} > 0\right)$.

Линейное накопление информации осуществляется, например, в процессах книгопечатания, кристаллизации, редупликации хромосом и т. п. Примером прогрессивного накопления информации могут служить все эволюционные процессы, осуществляющиеся не только в живой, но и в неорганической природе, процессы создания новых научных теорий, технических конструкций или социальных структур. Из проведенного анализа следует, что **<u>прогрессивное накопление информации</u> не может осуществляться по детерминированной программе.** В основу таких процессов (равно как и их электронных моделей) должен быть непременно заложен вероятностный, стохастичный механизм.

Оценивая возможности моделирования биологических процессов с помощью детерминированных моделей, А. В. Беркс пришел к заключению: «Мы не сможем как следует разобраться в вероятностных автоматах до тех пор, пока не определим границы возможностей детерминированных автоматов, т. е. не будем знать, что они не в состоянии выполнять». При этом Беркс предполагает, что решать эту задачу следует эмпирическим путем. Однако, анализ детерминированных автоматов свидетельствует о том, что границы возможностей детерминированных автоматов могут быть установлены не только экспериментально, методом «проб и ошибок», но и путем логических заключений, базирующихся на общетеоретических положениях и накопленном опыте применения вероятностных и детерминированных программ для электронно-вычислительных машин. Нет нужды доказывать, насколько более предпочтительным является именно этот путь.

Если в процессе включения новых элементов в систему (т. е. с увеличением *N)* число признаков, по которым осуществляется отбор элементов, остается неизменным (т. е. сохраняется постоянный «алфавит»), то соотношение $\frac{\partial^2 I}{\partial N^2} > 0$ будет выполняться только в том случае, если рост числа элементов системы *N* сопровождается перераспределением (дифференцировкой) вероятностей элементов p_i и соответствующим уменьшением энтропии, определяемой с помощью функции $\sum_i p_i \log p_i$.

К анализу происходящего в ходе эволюции процесса дифференцировки вероятностей мы теперь и перейдем.

16.2. Пример соотношения случайных и детерминированных связей: язык

В предыдущем изложении мы старались подчеркнуть объективную роль стохастичных связей в процессах развития. Однако из этого вовсе не следует, что детерминированным связям принадлежит в этих процессах какая-то второстепенная роль. В диалектическом взаимодействии случайных и детерминированных связей роли их равноценны, перевес в сторону случайных связей приводит к нарушению целостности системы, перевес в сторону детерминированных связей уменьшает способность адаптации системы к условиям внешней среды. Мир, в котором стохастичные связи начали бы повсеместно вытеснять связи детерминированные, был бы ничуть не лучше, чем детерминированный лапласовский мир. В таком мире постепенно перемешались бы все видовые признаки: конь с петушиным гребнем мог бы родиться с такой же степенью вероятности, как и петух с конским хвостом. А в конечном итоге в таком мире должны вообще исчезнуть всякие признаки и различия, ибо с исчезновением детерминированных связей и при безраздельном господстве связей случайных должна наступить та самая тепловая смерть, при которой Вселенная превратилась бы в лишенный каких бы то ни было признаков «бульон» из хаотически движущихся частиц. К счастью, этого не происходит, поскольку, как мы отмечали, на «сосуществовании» случайных и детерминированных связей зиждется наш диалектический мир.

Различие между двумя видами связей подчеркивалось многими авторами. Однако эти типы связей рассматривались не в эволюционной динамике, а в состоянии статистического равновесия тех или иных сформировавшихся систем. В этом отношении А.Е. Седов пошел несколько дальше и рассмотрел динамический эволюционный процесс с помощью функции

$$\sum_i p_i \log p_i.$$

В качестве рабочей модели развивающейся системы выбран письменный текст. Выбор этот не случаен. Письменный текст есть наиболее изученная статистическими методами эволюционирующая система. В отличие от генетических кодов, глубоко спрятанных природой от исследователя, статистические взаимосвязи между элементами текста (буквами, словами, фразами) осуществляются буквально у нас на глазах.

Представим себе ретроспективно (и достаточно умозрительно) процесс развития языка как системы начиная с момента появления первых членораздельных звуков до наших дней. Как и всякий процесс развития, данный процесс характеризуется накоплением информации (увеличением набора грамматических и фонетических правил) и соответствующим уменьшением энтропии языка. Математически этот процесс выражается как тенденция к увеличению неравномерности распределения вероятностей букв (к дифференцировке значений p_i, входящих в формулу $H = -\sum_i p_i \log p_i$) (наглядной иллюстрацией процесса упорядочивания текста по мере увеличения дифференцировки вероятностей p_i являются «фразы», приведенные ранее. Каждая из трех фраз может служить моделью определенного этапа эволюции языка).

Увеличение набора правил, регламентирующих статистическую неопределенность, теория информации трактует как увеличение избыточности текста. Для инженера связи информация эта и в самом деле является избыточной: она не дает знакомому с грамматикой получателю новых сведений и потому лишь зря загружает предназначенный для передачи сигналов канал. Но это вовсе не значит, что данная информация не нужна для самого языка (точно так же информация о том, что на свете появился еще один петух, имеющий гребень и хвост не из волос, а из перьев, является избыточной для получателя этого сообщения, но отнюдь не излишней для генетического кода самого петуха).

Напротив, ***«избыточная» информация*** — это ***детерминированная информация, благодаря которой сохраняется структурная целостность любой развивающейся системы*** (в частности, языка). Заложенный в систему ***стохастический механизм («мутации») служит источником новой, непредсказуемой информации***, передаваемой данной системой (обычно после соответствующей фильтрации) в каналы связи других систем (получателю текста, в генетический код потомства, тканям развивающегося организма и т. п.).

Если считать, что все 32 буквы алфавита имеют равную вероятность, то подсчитанная по формуле

$$H = -\sum_i p_i \log p_i$$

неопределенность появления каждой буквы составит $H_{\max} = 5$ бит. Мы говорили, что статистические исследования русского языка показали: после учета всех возможных предсказаний (т. е. исключения избыточности) фактическая неопределенность каждой буквы русского текста ($H_о$) составляет в среднем примерно 1 бит и, таким образом, в среднем на каждую букву текста 4 бита из 5 бит будут избыточными. Другими словами, 80% (4/5) информации, содержащейся в письменных текстах, повторяются неизменно и лишь подтверждают незыблемость грамматических и фонетических правил.

Как было показано ранее, количество избыточной информации ($I_И$) служит одновременно и мерой сохраняемого системой детерминированного порядка, выражаемого через $H_П$, причем

$$H_п = H_{\max} - H_0 = I_и.$$

Для дальнейшего анализа соотношения случайных и детерминированных связей А.Е. Седов предложил ввести ***коэффициент стохастичности***:

$$G = H_0/H_п. \qquad (16.1)$$

Сопоставляя это выражение с используемым теорией информации выражением для коэффициента избыточности

$$R = 1 - \frac{H_0}{H_0 + H_п},$$

нетрудно установить, что введенный коэффициент G связан с коэффициентом избыточности R соотношением

$$G = \frac{1 - R}{R} \qquad (16.2)$$

(сопоставляя величину R с результатами анализа объективного содержания коэффициента избыточности R, отметим, что в данном случае рассматривается коэффициент R=$R_{\lim}$, отражающий объективную упорядоченность систем).

Как следует из вышесказанного, коэффициент стохастичности языка приблизительно равен G = 1/4. Это значит, что в данной системе информация является на 80% детерминированной, а на 20% стохастичной. Такое соотношение обеспечивает, с одной стороны, необходимую информативность, лабильность и гибкость, с другой — структурную устойчивость языка. Коэффициент G = 1/4 для языка, по-видимому, приближается к оптимальному, обеспечивающему своеобразный гомеостазис, в то время как любая из крайностей (абсолютная детерминированность или максимальная стохастичность) ведет к потере информативности языка. Данное соотношение возникло как результат эволюции языка, выражающейся в отмеченной выше дифференцировке значений вероятностей p_i. С точки зрения интересующего нас соотношения случайных и детерминированных связей процесс дифференцировки p_i есть не что иное, как постепенный переход от связей «чисто случайных» (когда все вероятности одинаковы, H_0=$H_{шах}$, $H_п$=0, а G бесконечно велико) к «почти детерминированным» связям, когда вероятность одних событий приближается к единице, а других почти исчезает (стремится к нулю). При этом увеличение одних вероятностей осуществляется за счет уменьшения других в силу условия:

$$\sum_i p_i = 1 \cdot$$

Условие

$$\sum_i p_i = 1$$

справедливо для случаев неизменного «алфавита» (числа признаков), так как с точки зрения теории вероятностей такой алфавит представляет полную группу событий.

Экстраполируя этот процесс, легко прийти к заключению, что пределом его является полная детерминированность, соответствующая максимальной дифференцировке значений p_i, когда вероятность одной из букв становится равной единице, а всех остальных — нулю: например, $P_А = 1$, $P_Б = P_В = \ldots = P_Я = 0$. Таким образом, пределом процесса развития языка является текст из одинаковых букв (например, из одних А). **Этот случай, которому соответствуют условия $H_o = 0$, $H_п = H_{max}$, G=0, есть не что иное, как <u>модель системы</u>, которая, достигнув <u>предела развития</u>, стала полностью <u>детерминированной</u> и вследствие этого утратила возможности вариаций, информативность, способность к дальнейшей эволюции, а следовательно, и жизнеспособность.** Единственным условием сохранения существования подобной детерминированной системы является ее узкая специализация. Так, например, текст из одних букв *А* имеет право на существование в том случае, если в букву *А* заложен однозначный сугубо специализированный смысл; например, сообщение *А* может обозначать, что груз прибыл в заранее обусловленное место.

Надо заметить, что и в этом случае полная детерминация не достигается в силу того, что остается некоторая вероятность отрицательного ответа (т. е. продления паузы вместо ожидаемого сигнала *А*). Учитывая малую вероятность дорожных происшествий, препятствующих своевременной доставке груза, мы приходим к выводу, что в данном случае текст вырождается в малоэнтропийный (т. е. «почти детерминированный») двузначный код (пауза — *А*). **Предел детерминации такого кода представляет собой регулярное чередование пауз и сигналов *А*. Это уже не что иное, как закодированный периодический процесс.**

Рассмотренный пример текста из одних букв *А* соответствует предельной дифференцировке значений вероятностей отдельных букв без учета межбуквенной корреляции. При учете же межбуквенных связей предельной детерминированности системы будут соответствовать повторяющиеся сочетания букв (например, ПАМ, если $p_{ПАМ} = 1$, а вероятность всех остальных сочетаний равна нулю). При учете связей более высоких порядков (между словами, фразами) пределом упорядоченности будет повторение одной фразы или набора стереотипных фраз. Общее свойство всех перечисленных предельных случаев заключается в полной предсказуемости сообщений, вытекающей из их повторяемости.

Эти выводы можно распространить на все «языки», которыми пользуется природа. В частности, **безусловный рефлекс — это не что иное, как сохраняемая нервной системой «стереотипная фраза», которой организм отвечает на определенное сочетание внешних сигналов.** Подобно тексту из одних букв *А,* биологический вид, наделенный сочетанием жестко детерминированных рефлексов, инстинктов и наследственных признаков, **способен сохранять жизнеспособность только в условиях неизменной среды**.

Все системы, которые в результате эволюции достигли полной детерминации, ***предельно адаптированны, но абсолютно неадаптивны*.**

В масштабах планетных систем абсолютно стабильной средой является космос. Именно поэтому в процессе космогонической эволюции распыленная в космосе (энтропийная) материя постепенно превращается в системы планет, детерминированных настолько жестко, что астрономия с полной гарантией может предречь, что 16 октября 2126 г. в Москве будет наблюдаться солнечное затмение.

Можно привести и другие примеры, подтверждающие, что в некоторых случаях в результате эволюции достигается высокая степень детерминации. И тем не менее утверждение, что **пределом любого развития является узкая специализация, обусловленная полной детерминацией системы, справедливо при одной существенной оговорке: рассматриваемый процесс развития осуществляется на неизменном структурном уровне.**

Действительно, при определении упомянутых величин $H_п$ $(I_и)$ и H_O письменных текстов учитываются связи между буквами и словами, но не рассматриваются связи между фразами, понятиями и т. д. Поэтому описанный нами процесс развития, пределом которого является текст из одних букв *А,* следует рассматривать лишь как частный случай — случай, в реальных условиях приводящий к образованию косных теоретических концепций, консервативных (бюрократических) социальных структур, тупиковых биологических эволюционных ветвей и т. п. В более общем случае процессы развития осуществляются по ступеням: **исчерпав резервы мутаций на заданном уровне, система может переходить на уровень более сложной структуры и продолжать развиваться на нем** (при переходе на более высокий структурный уровень вступает в действие закон прогрессивного увеличения разнообразия, рассмотренный ниже).

Помимо ограничения уровня развития в рассматриваемом процессе подразумевалась также неизменность алфавита и стационарность условий отбора (набор правил, сформировавшийся в начале процесса, сохраняет свою силу на всех последующих стадиях).

Все сказанное относительно ограничений вовсе не значит, что наша модель бесполезна. Она «работает», как и любая другая теоретическая концепция, для создания которой необходимо от каких-то реальных условий абстрагироваться полностью, а какие-то условия упростить. В конце концов даже самый сложный процесс развития может быть условно разбит на этапы, на которых с какой-то степенью приближения можно считать допустимым и условие неизменности «алфавита», и условие стационарности внешней среды. Несомненно, что при таком поэтапном подходе будет легче понять процесс в целом, распознать его механизм.

Одним из примеров, убеждающих в полезности такого подхода, может служить развитие человека. Сейчас общепризнано, что биологическое развитие человека фактически прекратилось. Оно перешло на более высокий уровень: развитие мысли и социальных систем. Сам процесс развития мысли может служить еще одним примером того, как, достигнув детерминации на заданном уровне, система продолжает развитие на уровне более сложном. Освоение алгебры для школьника является творческим процессом, но для математика элементарная алгебра — это в основном набор стандартных (детерминированных) правил, служащих вспомогательным инструментом при творческом анализе сложной взаимосвязи изучаемых структур. При этом мысль оперирует с понятиями более «высокого» структурного уровня (функции, функционалы, множества и т. п.), а «школьные» алгебраические представления присутствуют в этих понятиях и оперировании ими как «составные элементы».

Аналогичная закономерность характерна для процесса биологической эволюции: когда клеточные структуры достигли определенной степени специализации, они перестали изменяться, а основой дальнейшей эволюции стал более высокий структурный уровень — межклеточная взаимосвязь. В многоклеточных организмах узкая специализация и детерминированность функций таких органов, как сердце или почки, сочетается с широкой многозначностью функций мозга, нервной системы и механизма наследственности. При этом только благодаря деятельности недетерминированной регулирующей системы, обеспечивающей независимость целого ряда параметров (температуры, давления, состава крови и др.) от условий *внешней* среды, создается та самая стабильная *внутренняя* среда, в которой могут нормально функционировать детерминированные и узкоспециализированные органы (для которых эта среда является уже внешней).

Распространяя эти примеры на другие виды систем, можно сделать следующий вывод: ***в сложной многофункциональной системе узкая специализация некоторых ее функций и органов может сосуществовать с недетерминированной гибкостью других ее органов и частей.***

Возвращаясь к языку, это общее положение можно проиллюстрировать наглядным примером: многозначность, информативность и гибкость языка в целом безболезненно уживаются с лаконизмом и однозначностью многих специализированных языков (язык инструкций, уставов, финансовых и экономических отчетов и т. п.).

Предел развития (полная детерминированность) может быть достигнут на любом структурном уровне эволюционирующих систем. Повторение одинаковых букв или слов, стереотипных фраз, мыслей, концепций — все это примеры систем с низкой энтропией, для которых вероятность одной из комбинаций близка к единице, а всех остальных — к нулю. При этом можно отметить общую закономерность: если любой из структурных уровней полностью детерминирован, то возможность перехода системы на более высокий уровень полностью исключена. В самом деле, из одинаковых букв невозможно составить слово, из одинаковых слов — фразу, из одинаковых фраз — концепцию и т. д.

Стереотипность детерминированной системы (т. е. отсутствие вариаций в комбинациях ее элементов) является также причиной, исключающей ее адаптацию к условиям внешней среды.

Несмотря на уменьшение гибкости, детерминированная система может иметь преимущества по сравнению с «многозначной» системой, обладающей большим количеством случайных связей. Эти преимущества проявляются в случае узкого диапазона изменения внешних условий, к которым адаптирована система; заключаются они в быстроте реакции и однозначности решений специализированных задач. Действительно, вместо фразы «груз прибыл по назначению» быстрее и проще передавать условный сигнал *А*. Однако стоит лишь измениться условиям (например, изменить место назначения груза), и детерминированная система окажется непригодной: текстом из одних букв *А* сообщить об изменении места доставки нельзя. При отклонении от заданного диапазона условий более жизнеспособной и адаптивной безусловно оказывается система, обладающая большим количеством гибких вероятностных связей. Конкретным подтверждением этого положения могут служить детерминированные и узкоспециализированные инстинкты животных, часто заставляющие их обладателей при изменении внешних условий действовать «невпопад» (так, например, после удаления матки из пчелиного улья пчелы продолжают носить пищу, оберегать несуществующую матку от опасностей и т. п.).

Этот пример является частным проявлением общей закономерности, которую В. И. Свидерский и Р. А. Зобов характеризуют так: «Уже очень давно было замечено, что возрастающее приспособление организмов к среде приводит к падению их жизнеспособности — ухудшению конституции, падению выносливости, темпов развития, плодовитости, снижению скоростей биохимических реакций и т. п.». Такова цена, которую приходится платить организмам, обладающим высокой степенью детерминации и узкой специализацией функций, достигнутой путем адаптации к условиям стабильной среды.

В узком диапазоне условий общее преимущество детерминированной программы заключается в том, что такая программа более «экономична», поскольку она избавляет от необходимости совершения большого числа ошибочных проб. Общую направленность процессов развития к детерминации следует расценивать как стремление природы избавиться от расточительных проб и ошибок во всех случаях, когда удается обходиться без них. Однако при воздействиях сложной, изменчивой внешней среды чрезмерная детерминированность не приносит системе ничего, кроме вреда.

Узкий специалист способен в частных вопросах разбираться быстрее и эффективнее, чем человек с большой эрудицией и широким кругозором; в то же время он проявляет беспомощность при столкновении со смежными областями или при вторжении в его узкую область научных или технических новшеств (этот вывод подтверждается результатами конкретных информационных оценок. Например, установлено, что опытная стенографистка способна воспринять и зафиксировать до 50 бит информации в секунду, в то время как человек в среднем воспринимает около 5—6 бит. Что же касается глубины восприятия информации, ее рационального отбора и эффективного использования, то тут в выигрыше оказывается не стенографистка, а специалист в данной области).

В противоположность этому знание смежных областей часто является источником случайных связей (ассоциаций), лежащих в основе многих открытий, возникающих на стыке наук (заметим, кстати, что сама теория информации служит тому примером: при решении задачи создания оптимальных систем связи у К. Шеннона возникла ассоциация с выведенной Л. Больцманом формулой физической энтропии, давшая затем неожиданные и удивительные плоды).

В данном случае оптимальным, по-видимому, является сочетание глубокой специализации в одной узкой области знаний с широким охватом основных сведений из ряда других областей. Каждый мыслящий человек на определенном этапе жизни индивидуально решает эту проблему (ее можно назвать «проблемой дилетантизма»), выбирая для себя оптимум распределения времени для совершенствования в узкой специальности и для удовлетворения, более широкого круга интересов (смежные области, искусство, природа, общение с людьми и т. п.).

К соотношению G_{opt}, близкому к оптимальному, приводит эволюция различных систем. Как мы отмечали, именно такое соотношение степени стохастичности и детерминированности обеспечивает функциональную гибкость устного и письменного языка. Благодаря тому, что в процессе эволюции язык не достиг предела детерминизма, а сохранил стохастические свойства, он не утратил своей многозначности, делающей его универсальным средством человеческого общения.

Всякая эволюционирующая система способна сама находить оптимум соотношения детерминированности и стохастичности (точнее, приближаться к нему), поскольку эволюция является результатом диалектического единства и «борьбы» противоположных тенденций — стремления к детерминации и ко всякого рода «мутациям». Благодаря «мутациям» в языке, как и в гене, возникают и накапливаются новые полезные признаки: литературные неологизмы, новые научные термины, слова, ассимилируемые из иностранных языков, и т. п.

Используя свой личный опыт, каждый из нас имеет возможность проследить своеобразный процесс «детерминизации языка». Как показал К. И. Чуковский, наиболее творческое отношение к образованию слов и построению фраз наблюдается в период освоения разговорной речи в возрасте от двух до пяти лет. Действуя методом проб и ошибок, ребенок образует всевозможные, часто совсем нелепые слова и сочетания слов. При этом порой возникают и весьма любопытные словообразования, удовлетворяющие формальным правилам разговорного языка. Вот некоторые примеры: песок *песучий;* обезьяны *уклюжие; мазелин, вертилятор, копатка, строганок, колоток; отпомнил, распонял; вык, вык и привык; отмухиваться* (отмахиваться от мух); *луксус* (лук в уксусе). Раз на ногах — *ногти,* то на руках — *рукти;* раз гуси ходят *гуськом,* значит утки — *утьком.* Множество других тщательно собранных и систематизированных примеров можно найти в упомянутой нами замечательной книге К. И. Чуковского. По окончании периода освоения разговорной речи человек перестает заниматься словообразованием и в дальнейшем пополняет свой речевой арсенал только путем заимствования общепринятых (детерминированных) выражений и слов.

В более позднем возрасте (примерно от 10 до 25 — 30 лет) происходит аналогичный процесс «детерминизации» представлений, понятий и взглядов. Исключением из этого правила являются ученые и художники слова, которые сохраняют нестандартность мышления или творческое отношение к слову на всю жизнь. Именно благодаря этому свойству отдельных незаурядных умов появляются на свете такие неожиданные оригинальные теории, как теория эволюции Дарвина или теория относительности Эйнштейна. Тем же свойством творческой личности обусловлено создание выдающихся произведений искусства, поражающих оригинальностью взглядов их авторов и неожиданностью найденных ими выразительных средств (в частности, таких образных неологизмов, как «стушеваться» — Ф. Достоевский, «обиностраниться» — Н. Гоголь, «в смокинг вштопорен», «отношение плевое» — В. Маяковский, «поступь надвьюжная» — А. Блок). Как в том, так и в другом случае — то есть и при абстрактном и при образном мышлении — залогом продуктивности творчества является рассмотренное нами условие $G = G_{\text{opt}}$.

Чтобы показать, что введенный коэффициент $G = H_o/H_п$ позволяет, хотя бы в принципе, производить сравнительную оценку художественных образов, рассмотрим крайние случаи соотношений случайных и детерминированных связей на уровне слов. Например, при поэтическом творчестве случаю G=0, H_o=0, $H_п$=$H_{\max}$ соответствуют литературные штампы: «синее небо», «ясный взгляд» и т. п. Другому крайнему случаю (G=∞, $H_o = H_{\max}$, $H_п = 0$) соответствует равная вероятность любых сочетаний слов. Это значит, что на каждое осмысленное словосочетание приходятся миллионы либо бес-смысленных («умный камень», «вкусный деготь» и т. п.), либо вовсе не согласованных грамматически сочетаний слов (вроде: «громко гвоздь бежала над светлое»). И только при оптимальных значениях параметра G (т. е. при $G \cong G_{\text{opt}}$) возникают яркие, творческие литературные образы вроде «флейты-позвоночника» или «страны березового ситца». Обобщая рассмотренные примеры, можно сделать вывод, что крайнему абстракционизму или формализму в искусстве соответствует $G \to \infty$, а полному следованию традиционным формам $G \to 0$. Объективным решением спора между представителями двух этих крайних направлений являются художественные произведения, соответствующие G_{opt}.

Убедительной иллюстрацией рассмотренных закономерностей может служить существующее в некоторых странах направление сценического искусства, получившее название «театр абсурда». Основным приемом драматургии театра абсурда является такое построение действия и диалога, при котором каждый из персонажей, не вникая в смысл высказываний партнеров, говорит о своем. В принятых нами обозначениях подобная ситуация характеризуется увеличением коэффициента стохастичности G на уровне слов и фраз. Именно это условие порождает абсурдность коллизий, характерных для сюжетов театра абсурда. В некоторых случаях такой прием может служить средством сильных художественных воздействий и даже философских обобщений. Его можно расценивать как художественную гиперболизированную форму постановки проблемы коммуникабельности человека в наш сложный и динамичный век. Однако злоупотребление этим сценическим приемом, выражаемое условием $G\to\infty$, приводит к тому, что зритель перестает понимать смысл происходящего на сцене и, не улавливая взаимосвязи событий, теряет к нему интерес. Аналогичная ситуация возникает в изобразительном искусстве в случае, если художник, стремящийся к абстракционизму, перестает доносить до зрителя взаимосвязь избранных им изобразительных средств, красок и форм.

16.3. Общие свойства эволюционных процессов.

На основе рассмотренных закономерностей и с учетом введенных выше обозначений можно дать несколько дополнительных важных определений.

Прогрессивным процессом следует считать процесс, соответствующий условию $G \to G_{opt}$, независимо от того, растет ли при этом число случайных связей или увеличивается детерминированность системы. К числу прогрессивных процессов с увеличением детерминированных связей относятся все процессы накопления информации. Из проделанного нами анализа следует, что эти процессы остаются прогрессивными лишь до тех пор, пока выполняется условие $G > G_{opt}$. В случае, когда $G < G_{opt}$, прогрессивным является процесс увеличения случайных связей, происходящий за счет разрушения связей детерминированных (при этом $G = H_o/H_п$ возрастает). Примером таких процессов могут служить революционные взрывы, разрушающие детерминированные связи социальных структур.

Процесс, соответствующий условию $G\to\infty$, можно назвать *деградацией.* Семантика этого термина полностью соответствует сущности данного процесса: исчезновение различий (градаций) при достижении состояния, соответствующего условию $H_0 = H_{max}$

(указанное соответствие возникло не случайно. Первоначально под термином «деградация» подразумевалось исчезновение градаций макроскопических признаков в структуре систем, но в ходе дальнейших исследований стало очевидным, что деградация макроструктуры обусловлена деградацией микроструктуры, т. е. условием $H_0 = H_{max}$).

Процесс, описываемый условием $G \to 0$ (детерминированный текст и т. п.), лучше всего передать термином *вырождение.*

Заметим, что обычно термины «деградация» и «вырождение» использовались как синонимы. Теперь выясняется, что их целесообразно применять для обозначения двух полюсных состояний эволюционирующих систем. Вместе с тем смысл их не так уж различен: система в стадии вырождения теряет способность адаптации и потому легко деградирует (разрушается) при изменениях условий среды. Если от стадии деградации к стадии вырождения ведет долгий путь эволюции, то путь от вырождения к деградации представляет собой, как правило, лишь короткий скачок (рис. 16.2).

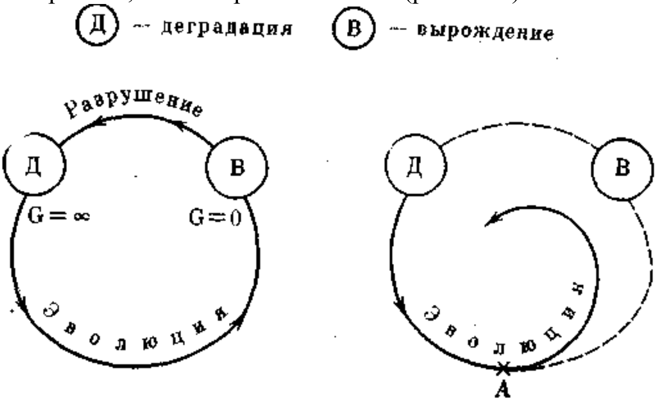


Рис. 16.1. Рис. 16.2

Для поддержания всеобщего равновесия природа «предусмотрела» и возможность обратных скачков. Такие скачки осуществляются в уже рассмотренном нами превращении тепловой энергии в энергию излучений. Это превращение происходит путем квантования энергии, в результате которого энтропийное тепловое движение приобретает определенную степень детерминации, выражающуюся в волновом характере излучений. (Предельным случаем является монохроматическое излучение, представляющее собой жестко детерминированный периодический процесс.) В обозначениях рис. 16.1 данному процессу соответствует скачок из состояния Д в состояние В.

Как следует из предыдущего анализа, цикл, изображенный на рис. 16.1, может осуществляться лишь на неизменном структурном уровне эволюционирующих систем. При переходе на более высокий уровень «порочный круг», изображенный на рис. 16.1, разрывается (точка А на рис. 16.2) и кривая развития превращается в эволюционную спираль.

Представленная на рис. 16.2 схема может служить наглядным опровержением метафизических концепций, стремившихся доказать, что миру в конечном счете должна быть присуща одна из крайностей: либо провозглашенная Лапласом полная детерминация (условно обозначенная на рис. 16.2 буквой В), либо предрекаемая многими учеными (включая Н. Винера тепловая смерть (буква Д на рис. 16.2).

В реальном мире системы, достигшие этих крайних состояний, являются не правилом, а исключением: полной детерминации (В) эволюционным путем достигают едва ли не одни лишь планетные системы, а полной стохастичности (Д) — находящийся в термодинамическом равновесии, да притом еще и идеальный, газ (в реальных газах приходится учитывать взаимодействие (т. е. корреляцию траекторий) молекул, приводящее к уменьшению энтропии. В идеальном газе Максвелла реальные взаимодействия молекул заменены моделью взаимодействия упругих шаров).

Большинство же реальных систем достигает гомеостазиса, находясь в промежуточном состоянии между двумя полюсами, обозначенными на рис. 16.1 и 16.2 через В и Д. Это состояние, при котором в системе сохраняются и детерминированные, и стохастичные связи, характеризуется введенным Е.А.Седовым коэффициентом стохастичности *G*.

Теперь становится понятным, почему в возникшем еще в античные времена и длящемся 2500 лет споре о соотношении необходимого и случайного чаши весов попеременно склонялись то в пользу В, то в пользу Д. В этом проявлялась диалектика самого процесса познания мира. Чтобы выявить тот «компромисс» между необходимостью и случайностью, которого достигают самоорганизующиеся системы, недостаточно было не только самих понятий «необходимого» и «случайного», но даже и их вероятностных количественных мер,— нужно было сначала выработать такие интегральные характеристики порядка и стохастичности, как меры количества информации и энтропии. Лишь благодаря им становится возможной оценка соотношения случайного и необходимого — и в реальных гомеостатических и эволюционных процессах, и в их теоретическом отражении в науке.

Поскольку все существующие в мире естественно развивающиеся системы исследуются в течение ограниченных временных интервалов, а процессы биологической эволюции длятся сотни тысяч, а порой и миллиарды дет, мы наблюдаем, как правило, не процесс эволюции, а лишь его результат. Однако для уяснения диалектики случайных и детерминированных связей необходимо выявить динамику эволюционных процессов. Эффективными инструментами такого анализа опять-таки являются используемая теорией информации функция $\sum_i p_i \log p_i$ и вычисляемый с ее помощью коэффициент стохастичности $G = H_0/H_п$. При этом выражаемые через H_o и $H_п$ случайные и детерминированные связи предстают как неразрывно связанные противоположные сущности, о которых говорят: «...ум человека не должен брать эти противоположности за мертвые, застывшие, а за живые, условные, подвижные, превращающиеся одна в другую».

Происходящее в ходе эволюционного процесса изменение отношения $H_0/H_п$, выражаемое уменьшением параметра *G*, отражает именно эту подвижность, т. е. процесс превращения одной противоположности (стохастичности), в другую (детерминированность).

Представим себе достигшую определенной стадии развития систему, у которой в данный момент соотношение случайных и детерминированных связей составляет $G=H_0/H_п$. Как следует из предыдущих рассмотрений, величина $H_п$ характеризует степень упорядоченности существующего в системе движения и равна количеству информации, содержащейся в *наборе правил,* регламентирующих этот порядок (грамматических и фонетических правил, наследственных признаков и т. п.). Величина H_o характеризует количество «импровизаций» в пределах, дозволенных всеми правилами, которым соответствует величина $H_п$. Появившиеся в результате мутаций новое слово или новый наследственный признак в генетическом коде должны, с **одной стороны, удовлетворять всем установленным ранее правилам, а с другой — отличаться от всех возникших ранее «слов».**

На начальном этапе развития, когда H_o велико, а $H_п$ мало (т. е. при больших *G*), набор «правил» настолько мал, что удовлетворить им весьма несложно. Поэтому число комбинаций «букв» заданного «алфавита» (или, пользуясь термином М. Планка, число комплексий), удовлетворяющих всем правилам, достаточно велико. Это значит, что для начальной стадии развития характерно ускоренное образование разновидностей несложных форм. Именно так протекает проанализированный К. И. Чуковским процесс языкового творчества у детей. Вопрос, сохранит ли система новые признаки того или иного мутанта на всех этапах развития, решается взаимодействием «система

— среда». В числе разновидностей, возникающих в ходе взаимодействия, всегда есть и такие, которые несут в себе не только признаки, предусмотренные набором ранее сформировавшихся правил, но и ряд новых. Это и есть те новые слова или биологические особи, которым вместе со всеми иными мутантами предстоит быть признанными или отвергнутыми в ходе отбора.

Подводя итог сказанному, мы приходим к выводу, что на начальной стадии за счет больших значений H_o формообразование осуществляется быстро и темпы развития достаточно велики. В ходе развития H_o падает, а $H_п$ растет. Следует подчеркнуть, что при неизменном числе разновидностей элементов системы (т. е. при неизменном «алфавите») выполняется условие: $H_o + H_п = H_{max}$ = const (например, в случае алфавита из 32 букв письменный текст независимо от степени его детерминации (т. е. при любых $H_п$ и H_o) удовлетворяет условию: $H_o + H_п = H_{max} = 5$ бит = const). Это значит, что увеличение $H_п$ приводит к уменьшению $H_о = H_{max} — H_п$, а это способствует еще более быстрому уменьшению отношения $G = H_0/H_п$, в котором числитель падает, а знаменатель растет. По мере уменьшения H_o мутаций становится меньше, а благодаря увеличению $H_п$ все меньшее число мутантов удовлетворяет набору всех «правил». Следовательно, с **уменьшением параметра *G* все меньше и меньше образуется разновидностей новых слов или наследственных признаков, <u>темп развития падает</u>.** Пределом этой тенденции является рассмотренная выше детерминированная система, в которой мутации и формообразование прекратились совсем (текст из одних букв *А,* планетная система и т. п.)

До сих пор речь шла о диалектическом взаимодействии случайных и детерминированных связей ***внутри*** эволюционирующей системы. Теперь попытаемся с тех же позиций рассмотреть ***взаимодействие между системой и средой***. Для облегчения рассуждений рассмотрим процесс взаимодействия «система — среда» на упрощенной модели, в которой один рассматриваемый признак приспосабливается к одному из регулирующих параметров среды, причем распределение вероятностей признака и регулирующего параметра описывается нормальным законом.

Для изменяющегося по случайному закону (мутирующего) признака среда является своего рода «статистическим фильтром», «полоса пропускания» которого определяется дисперсией кривой распределения вероятностей значений регулирующего параметра среды ($\sigma_{рп}$) (эту «полосу пропускания» В. А. Геодакян называет «шириной экологической ниши по данному фактору среды»). В предельном случае, когда $\sigma_{рп} = 0$, возникает отбор признака, который А. Л.

Тахтаджян называет «матричным». При таком отборе среда как бы штампует идентичные признаки систем.

Другой предельный случай выражается условием, когда $\sigma_{рп}$ примет «бесконечно большое» значение. В этом случае «полоса фильтра» неограниченна, следовательно, никакого отбора не осуществляется, система с любой дисперсией по данному признаку может находиться в равновесии со средой. Все промежуточные случаи, удовлетворяющие условию $0<\sigma_{рп}<\infty$, соответствуют процессу ***эмергентного*** (созидательного) отбора мутирующих признаков систем.

Если пределы мутаций признака ($\sigma_{рп}$) превышают дисперсию регулирующего параметра ($\sigma_{рп}$), то часть мутантов будет не соответствовать указанному параметру, и, следовательно, произойдет их отбраковка (естественный отбор). Противоположное условие ($\sigma_{мп}<\sigma_{рп}$) соответствует случаю, когда признак системы настолько детерминирован, что возможные пределы его мутаций не обеспечивают необходимого соответствия во всем диапазоне изменений регулирующего параметра. **Оптимальным условием адаптации является условие $\sigma_{мп}=\sigma_{рп}$.**

Для иллюстрации этих положений рассмотрим пример. Научная школа, признающая лишь те идеи, которые соответствуют определенным традициям, является «узкополосным фильтром», осуществляющим матричный (точнее, тяготеющий к матричному) отбор. Система установившихся взглядов — это и есть та самая «матрица», которая в данном случае детерминирует отбор. Для творческого (эмергентного) отбора научная среда должна быть достаточно восприимчива и лабильна, чтобы захотеть и суметь вникнуть в суть даже совсем непохожих на прежние «сумасшедших» идей.

Теперь рассмотрим другой крайний случай: научный кворум готов принять самые неожиданные идеи, но мысли научных работников настолько детерминированы (скажем, влиянием прежней научной школы), что ни один из них не способен предложить каких-либо свежих идей. В этом случае диапазон предлагаемого научным сотрудником творческого поиска уже тех требований, которые предлагает ему научная среда. Среда отвергает такого сотрудника как неспособного. Точно так же **естественная среда отвергает (уничтожает) слишком детерминированный биологический вид, если в течение нескольких поколений он не способен путем мутаций «нащупать» все признаки, удовлетворяющие диапазону изменений условий среды.**

Если пределы мутаций признаков (или научных идей) совпадают с диапазоном критериев их отбора, то процесс осуществляется без излишних потерь.

Этому случаю соответствует условие $\sigma_{мп}=\sigma_{рп}$. Очевидно, что дополнительным условием продуктивности творческого процесса является оптимальный выбор величины $\sigma_{рп}$, т. е. установление разумных пределов отбора неожиданных новых идей.

Все приведенные примеры свидетельствуют о том, что в результате отбора распределение вероятностей ***мутирующих признаков*** в конечном счете «вписывается» в кривую распределения вероятностей ***внешних условий***. Поскольку определяющие значения коэффициента G величины H_o и $H_п$ так же, как и дисперсии $\sigma_{рп}$ и $\sigma_{мп}$, являются функциями распределения вероятностей p_i, можно предполагать, что ***общим условием оптимальной адаптации (гомеостазиса) системы является равенство коэффициентов стохастичности системы и среды,*** то есть условие:

$$G_{\text{системы}} = G_{\text{среды}} = G_{\text{opt}}.$$

Справедливость данного соотношения подтверждается анализом процессов адаптации, которые в самом общем виде могут быть также представлены как процессы дифференцировки значений p_i функции $\sum_i p_i \log p_i$.

Представим себе систему, обладающую возможностью выбирать один из n возможных ответов (реакций) на воздействия внешней среды. В исходном состоянии система способна отвечать на любые воздействия лишь выбранной наугад реакцией i $(i = 1, 2, \ldots, n)$. Это значит, что вероятности всех ответов системы p_i будут равны. Как было показано в предыдущем пункте, этому случаю соответствуют условия :

$$H_o = H_{\max}, H_п = 0, G_{\text{системы}} =\infty.$$

В результате обучения система приобретает способность дифференцировать ответы, варьируя их в зависимости от характера воздействий среды (в более общем случае — от сочетания этих воздействий). Примером подобного обучения может служить выработка условных рефлексов, позволяющая животным почти безошибочно выполнять устные распоряжения дрессировщика, осуществлять выбор одного из многих предметов и т. п. Как было установлено П. К. Анохиным, многие реакции имеют упреждающий характер. Это значит, что в результате обучения система приобретает способность на основании испытанных воздействий выявить присущую этим воздействиям корреляцию, «предвидеть» наиболее вероятные последующие воздействия и заранее давать соответствующий ответ. Таким образом, ***в ходе обучения устанавливаются вероятностные связи между воздействиями***

среды и ответными реакциями адаптирующихся систем, т. е. осуществляется дифференцировка вероятностей реакций p_i, причем наибольшей вероятности воздействий соответствует наибольшая вероятность ответов.

Рассмотренная закономерность процесса обучения может служить основой широких обобщений, поскольку ***процесс биологической эволюции включает в себя «обучение» поколений оптимальному поведению в условиях данной среды.*** Так же, как всякий процесс обучения, процесс биологической эволюции обусловлен накоплением информации, выражающимся математически как дифференцировка значений p_i функции $\sum_i p_i \log p_i$.

Обобщая рассмотренные закономерности на все случаи адаптации, мы приходим к выводу, что всякий ***процесс адаптации может быть представлен как процесс дифференцировки вероятностей различных реакций системы, соответствующих дифференцировке вероятностей различных воздействий среды.*** Указанное соответствие будет выражаться равенством значений $G_{(\text{системы})}$ и $G_{(\text{среды})}$, вычисляемых с помощью функции $\sum_i p_i \log p_i$. Частным случаем является условие $G_{(\text{системы})} = G_{(\text{среды})} = 0$, которое соответствует рассмотренным выше примерам полной детерминации системы в результате ее адаптации к условиям детерминированной стабильной среды.

Исследуемые закономерности процессов адаптации, приводящие к той или иной степени детерминации внешних и внутренних связей адаптирующихся систем, опровергают выдвинутую Н. И. Кобозевым гипотезу о существовании наряду с информацией (негэнтропией) некой «антиэнтропии» («энтропийного вакуума»), будто бы необходимой для возникновения живых организмов и осуществления детерминированных процессов мышления (формально-логических операций). Согласно предположению Н. И. Кобозева, «энтропийный вакуум» создается благодаря участию в этих процессах гипотических, обладающих ничтожно малой плотностью сверхлегких частиц.

Убедительным опровержением выдвинутой Н. И. Кобозевым концепции может служить практика создания электронных машин. В этих машинах удалось осуществить лишенные энтропии формально-логические операции, не прибегая к помощи «энтропийного вакуума» сверхлегких частиц (сказанное отнюдь не исключает возможности участия в психических явлениях и процессах мышления таких малоисследованных форм энергии, как нейтринные или гравитационные поля. Можно предполагать, что они играют важную роль при возбуждении, торможении и передаче информации (в частности, телепатической). Что же касается энтропии, то она, как мы убедились, может уменьшаться за счет воздействия и других видов полей (например, электромагнитного)).

При этом справедливо подчеркиваемая Н. И. Кобозевым энтропийность участвующих в этих процессах молекул и элементарных частиц (электронов, фотонов или магнонов) «преодолевается» путем перехода на макроскопический уровень, т. е. на уровень специальных электронных, оптических или магнитных устройств. На этом уровне жестко детерминированные связи обеспечиваются благодаря нелинейным (пороговым) свойствам указанных элементов, каждый из которых либо открыт (состояние 1 или «да»), либо закрыт (состояние 0 или «нет»). Переход элемента из одного состояния в другое соответствует элементарной операции отрицания двузначной логики, а совокупность операций, реализуемых такого рода элементами, позволяет получить жестко детерминированные формально-логические преобразования информации, т. е. «преодолеть» энтропийность участвующих в этих процессах микрочастиц. При этом, в противовес утверждениям Н. И. Кобозева, энтропия микродвижения компенсируется не «антиэнтропией сверхлегких частиц», а интегральным эффектом воздействия на пороговые устройства потоков обычных микрочастиц (электрических токов, магнитных потоков, световых лучей).

Пороговые устройства не реагируют на шумовые флюктуации (т. е. на энтропию), а изменяют свое состояние (срабатывают), когда суммарная энергия всех частиц потока превысит определенную наперед заданную величину. Нет никаких оснований утверждать, что на макроскопическом уровне коры головного мозга (т. е. на уровне нейронных сетей) без вмешательства «антиэнтропии сверхлегких частиц» принципиально неосуществимы те самые формально-логические операция, которые реализованы в современных ЭВМ.

Гораздо труднее предусмотреть в программах электронных машин «нарушения» формально-логических связей с целью получения эвристических решений. Для этого оказалась необходимой не «антиэнтропия», а обычная энтропия, вопрос только в том, каким образом и в какой пропорции ее «подмешать». В той же степени необходимой является энтропия и для процессов эвристического мышления, протекающих в человеческом мозге. А поскольку эвристическое мышление безусловно более сложное и ценное, чем формально-логическое, теряет смысл выдвигаемый Н. И. Кобозевым тезис о безинтропийности процессов мышления (в отличие от обладающих энтропией процессов, которые Н. И. Кобозев называет «процессами информации» и произвольно низводит в некий более низкий разряд (вместе с тем следует отметить, что высказанные в адрес работы Н. И. Кобозева критические замечания не умаляют многих ее достоинств, в частности оригинального и четкого анализа векторизации броуновских процессов в связи с соотношением связанной и свободной энергии, информации и энтропии).

Физический смысл вводимой Н. И. Кобозевым отрицательной энтропии остается неясным. Н. И. Кобозев определяет отрицательную энтропию как потенциальную возможность полной самоупорядоченности и самоорганизации систем. Проделанный нами анализ, напротив, показывает, что такой потенциальной возможностью обладают все адаптирующиеся системы с положительной энтропией, причем степень достигаемой ими детерминации и соответствующего уменьшения энтропии определяется степенью детерминации воздействующих на них факторов внешней среды.

Конкретной иллюстрацией общих закономерностей взаимодействий системы с внешней средой, приводящих к соотношению $G_{(\text{системы})}=G_{(\text{среды})}$, могут служить процессы эволюционных преобразований языка. Наибольшей степени детерминации язык достигает в периоды стабилизации социальных условий, тогда как в периоды общественных катаклизмов наблюдается наибольшая мобильность народного языка. Этот пример свидетельствует о том, что увеличение коэффициента стохастичности языка как системы является следствием увеличения коэффициента стохастичности среды.

Противоположным примером является процесс специализации языка в какой-либо узкой области применения. В применении к научному языку этот процесс М. Б. Воробьева описывает так: «Дело заключается не только во введении большого числа терминов, как каких-то новых образований, а в систематическом отборе из общелитературного языка определенных слов, в их частой повторяемости и в постепенном сокращении синонимии. Иными словами, характерной чертой развития лексики научного языка является прогрессирующее ограничение «общего словаря».

С позиций теории информации отбор и частая повторяемость определенных слов означает увеличение их вероятностей, а постепенное сокращение синонимии и прогрессирующее ограничение «общего» словаря — уменьшение вероятностей постепенно исключаемых из обращения слов. Таким образом, в процессе специализации происходит дифференцировка значений p_i, приводящая к уменьшению величины

$$\sum_i p_i \log p_i$$

и к детерминизации структуры языка на уровне слов. Как было показано п. 16.2, пределом этой тенденции является полное вырождение текста (повторение одинаковых слов).

Все рассмотренные примеры свидетельствуют о том, что всякий процесс эволюции — это процесс изменения соотношения случайных и детерминированных связей внутри системы, удобным критерием которого является коэффициент стохастичности *G*. Однако, используя этот коэффициент, необходимо учитывать, что его применение имеет ряд существенных ограничений. Одним из таких ограничений является конечность рассматриваемого набора признаков (алфавита) и вытекающее отсюда условие дискретного изменения признаков. Необходимость данного ограничения становится явной, если учесть, что значение H_{max}, определяемое как

$$H_{\max} = -\sum_i p_i \log p_i \text{ при } p_i = \frac{1}{n},$$

неограниченно возрастает с увеличением *n* (как $\log\limits_{n\to\infty} n$).

Пределом является непрерывное изменение признака, при котором величина H_o может быть вычислена по заданной кривой распределения вероятностей, с помощью формулы

$$H_0 = -\int\limits_{-\infty}^{+\infty} W(x) \log W(x)\, dx,$$

но величины H_{max} и $H_п = H_{max} — H_o$ становятся бесконечно большими, поэтому их вычисление теряет смысл (смысл условия $H_п \to\infty$ при $H_{max} \to\infty$ состоит лишь в том, что по сравнению с бесконечным беспорядком любой порядок бесконечно велик).

В этом случае критерием эволюционных преобразований могут служить только изменения непрерывных кривых распределения вероятностей и определяемых с их помощью величин H_o. Как отмечалось выше, рассмотренная нами модель взаимодействия эволюционирующего признака и регулирующего параметра среды является упрощенной. В реальных процессах эволюция каждого признака обусловлена не одним регулирующим параметром, а совокупностью параметров внешней среды. И наоборот, приспособление к изменению каждого из параметров среды обычно требует изменения целого комплекса признаков эволюционирующих систем.

Кроме того, во всех рассуждениях подразумевалась в неявной форме жесткая однозначная связь между изменениями признака и параметра среды. Тем самым частично дискредетировалась роль случайных связей в процессах адаптации систем. Между тем адаптация достигается только при том условии, что системе предоставлена определенная степень «свободы выбора» реакций на изменения параметров внешней среды. Это значит, что каждому значению управляющего параметра может соответствовать не одна, а несколько комбинаций признаков адаптирующихся систем. Для учета всех этих факторов помимо теории вероятностей необходимо использовать и аппарат теории множеств, и теорию графов, и топологию, и теорию игр. Однако, несмотря на усложнение аппарата и множество еще не решенных проблем, в любом случае будет полезной методология, основанная на анализе соотношения случайных и детерминированных связей путем анализа распределения вероятностей и вычисления значений H_o, $H_{\text{п}}$ и $G = H_0/H_{\text{п}}$.

Наряду с задачей теоретического исследования эволюционных моделей и их приближения к реальным процессам возникает необходимость эмпирического подтверждения выявленных на данном пути взаимосвязей и свойств. И в этом случае **главные трудности обусловлены сложностью и многозначностью связей между признаками системы и параметрами среды**. Комплексный характер взаимодействий реальных объектов практически исключает возможность вычленения в эксперименте какой-либо одной связи типа «признак системы — параметр среды». Поэтому приходится ограничиваться исследованием упрощенных моделей с помощью ЭВМ.

Определение значений G и G_{opt} в конкретных эволюционных процессах представляет не только теоретический, но и практический интерес. Зная эти параметры, можно производить количественную оценку своего рода «**эволюционного потенциала**»

$$П_{э} = G_{\text{opt}} - G,$$

показывающего, насколько далека система от оптимальной адаптации в заданных условиях внешней среды. Если предположить, что рассмотренное выше условие $G_{\text{opt}} = G_{\text{среды}}$ во всех случаях является справедливым, то задача определения G_{opt} всякой эволюционирующей системы сведется к определению коэффициента стохастичности тех или иных параметров внешней среды.

16.4. Общие статистические закономерности эволюции

Насколько важнным и актуальным является рассматриваемый в данном разделе вопрос о соотношении случайных и детерминированных связей, можно судить по высказываниям авторов ряда работ. Так, например, А. С. Кравец отмечает, что «проявление активности невозможно в условиях жесткой организации сложной системы, когда состояние любой подсистемы необходимым образом связывается с состоянием всех других подсистем. Функционирование такой жесткой системы возможно лишь по определенной изначальной программе, когда функции всех подсистем строго «расписаны» и взаимосвязаны. Любое отклонение от принятой программы ведет к разрушению системы. Автономная организация позволяет, наоборот, за счет активного выбора наилучшей стратегии максимально использовать ресурсы каждой подсистемы, быстрее решать поставленные задачи». «Мир,— заключает А. С. Кравец, — предстает как процесс постоянной борьбы противоположностей: изменчивости и устойчивости, связности и автономности, порядка и беспорядка, где каждая сторона может одерживать победы в некоторое время в одном месте и проигрывать в другое время и в другом месте».

Примерно в том же аспекте поднимается этот вопрос В. В. Налимовым и 3. М. Мульченко: «Вероятно, один из самых важных и интересных вопросов кибернетики можно сформулировать так: ***каким должно быть соотношение между самоорганизацией и централизованным управлением в большой системе для того, чтобы она была жизнеспособна.*** С этих позиций интересно было бы проанализировать существующие сейчас большие системы — технические, биологические, социальные».

Отмечая одну из основных тенденций научного мировоззрения, определяемую термином «системная ориентация», Э. Г. Юдин приходит к выводу, что «для науки в целом характерен переход от представлений о структуре как о том, что однозначно детерминирует, предписывает тот или иной ход событий, к представлениям о ней как о совокупности ограничений, накладываемых на «степени свободы» отдельных элементов системы,—ограничений, возникающих вследствие организованности элементов в рамках целого».

Сопоставив эти высказывания, нетрудно прийти к заключению, что речь здесь идет не о чем ином, как о *соотношении случайных и детерминированных связей*. В самом деле, *«автономия», или определенная «степень свободы», элементов системы может быть обеспечена только наличием <u>вероятностных связей</u>, исключающих* *жесткую детерминацию. И наоборот, жесткая детерминация системы исключает автономию ее элементов.* Так, например, Ю. В. Сачков определяет жестко детерминированную систему как систему, в которой «поведение любого объекта... определяется поведением всех остальных объектов, и притом взаимно однозначным образом... Схема жесткой детерминации отрицает какую-либо автономность в поведении элементов в рамках системы».

В несколько иных формулировках тот же вопрос о соотношениях случайности и детерминации в связи с проблемой возникновения жизни затрагивает Г. Патти. Он предлагает «отнести проблему возникновения жизни в мир физических законов и поместить ее между макроскопическими механическими системами, где возможна полная спецификация сил и динамических переменных, и статистическими системами с характерной для них *неполной спецификацией состояний».*

Отмечая, что случайные (вероятностные) связи обеспечивают относительную независимость элементов и частей (подсистем), Ю. В. Сачков приходит к заключению: «Где нет этой независимости подсистем, где все жестко определено, там по существу нет и самого управления в современной постановке этой проблемы, точнее, там может идти речь лишь о простейшем виде управления, основанном на культе приказа и его слепого исполнения». Поэтому «наиболее последовательными и математически разработанными теориями, учитывающими гибкость внутренних связей соответствующих объектов исследования, являются те, которые в математическом отношении опираются на теорию вероятностей». «Чтобы сложная система могла успешно и целенаправленно функционировать, между ее компонентами должна быть согласованность, включающая в значительной мере и жесткость связей. Вместе с тем компоненты систем обладают определенной независимостью (автономностью), которая, так сказать, по самому определению отрицает жесткость связей. Принцип структурной организации сложных систем, включающий строгую зависимость между элементами и их независимость, более широк и общ, чем принцип, основывающийся на одной лишь простой однозначной зависимости. Раскрытие взаимоотношения зависимости и независимости (автономности) лежит в основе познания сложных систем. Точнее говоря, зависимости (связи) не есть нечто такое, что может просто быть или не быть. Основной факт состоит в том, что они имеют внутренние градации по своей интенсивности, обладают большей или меньшей «степенью наличности», а раскрытие диалектики взаимосвязей реального мира включает в себя и опирается на анализ взаимопроникновения предельных (противоположных, взаимоисключающих) случаев». К сказанному остается лишь добавить, что **«степень наличности» связей и «градации» их «интенсивности» определяются значениями соответствующих условных вероятностей, а анализ «взаимопроникновения противоположных, взаимоисключающих случаев», как было показано в предыдущих пунктах, может производиться с помощью функции**

$$\sum_i p_i \log p_i.$$

Важность анализа соотношений случайных и детерминированных связей неоднократно подчеркивал Н. Винер. В частности, он писал: «Мы должны рассматривать организацию как нечто обладающее связью между отдельными организованными частями, причем взаимосвязь эта не единообразна. Связи между одними внутренними частями должны играть более важную роль, чем между другими,— иными словами, связи внутри организации не должны быть абсолютно устойчивыми, чтобы строгая определенность одних ее частей не исключала возможность изменения каких-то других. Эти изменения, различные в различных случаях, неизбежно носят статистический характер, и поэтому только статистическая теория обладает достаточной гибкостью, чтобы в своих рамках придать понятию организации разумный смысл». Непосредственную связь такого подхода с методами диалектики подчеркивает Ю. В. Сачков: «Весьма существенно, что в развитии общего учения о сложных управляющих системах... важнейшее значение приобрели проблемы диалектики, и прежде всего такие, как взаимопроникновение жестко- и аморфно-пластичного начал в структуре материальных систем, субординация и координация, широкая автономность систем и гармония целого, сохранение и коренное обновление...».

При сопоставлении всех приведенных высказываний обращает на себя внимание, с одной стороны, общность взглядов, а с другой — терминологический разнобой. Например, присущие сложным системам ***вероятностные связи*** именуются и «степенями свободы отдельных элементов» (Э. Г. Юдин), и «основой самоорганизации» (З. М. Мульченко и В. В. Налимов), и «автономностью организации» (А. С. Кравец и Ю. В. Сачков), и «неполной спецификацией состояний» (Г. Патти), и «возможностью изменений» (Н. Винер), и «аморфно-пластичным началом» (Ю. В. Сачков). Не менее разнообразно определяются и ***детерминированные связи***: «устойчивость, определенность» (Н. Винер), «жесткая организация» (А. С. Кравец), «централизованное управление» (З. М. Мульченко и В. В. Налимов), «однозначная детерминированность», «ограничение вследствие организованности элементов в рамках целого» (Э. Г. Юдин), «фиксированные ограничения разнообразия» (А. Л. Тахтаджян), «полная спецификация состояний» (Г. Патти), «гармония целого» (Ю. В. Сачков).

Сопоставление используемой различными авторами терминологии представляет для нас двоякий интерес. Во-первых, оно является наглядной иллюстрацией начальной стадии развития научных концепций, когда согласно рассмотренным в предыдущем пункте закономерностям $H_о$ велико, а $H_п$ мало, т. е. той стадии, когда всяческие мутации (в том числе и терминологические) возникают особенно интенсивно. Во-вторых, сопоставление приведенных формулировок с анализом, проведенным в предыдущем пункте, показывает, что применение таких емких и точных терминов, как ***детерминированность и стохастичность, и учет их соотношения с помощью коэффициента стохастичности G позволяют существенным образом упорядочить происходящее в данном направлении движение научной мысли, т. е. согласно введенным нами обозначениям, увеличить $H_п$ за счет уменьшения $H_о$.***

Следует подчеркнуть и то обстоятельство, что все приведенные выше высказывания характеризуют соотношение случайных и детерминированных связей в сформировавшихся системах с заданной неизменной структурой. Мы же пытаемся проследить эволюционную динамику структуры систем, выражающуюся в изменении соотношения случайных и детерминированных связей, т. е. в изменении введенного А.Е. Седовым параметра $G = H_о/H_п$. ***Простая связь параметра G с используемым теорией информации коэффициентом избыточности R позволяет применить накопленный опыт определения избыточности различных систем с целью оценки их эволюционных (созидательных, эмергентных) свойств.***

Попытку анализа процессов биологической эволюции с кибернетических позиций предпринял в своей научно-публицистической монографии «Сумма технологии» Станислав Лем. В частности, он писал: «У растения, бактерии или насекомого, как «гомеостатов первой степени», реакции на изменения среды заложены с момента рождения. Применяя язык кибернетики, можно сказать, что эти системы (особи) заранее «запрограммированы» по всем тем возможным изменениям среды, к которым они должны приспособляться для сохранения своей жизни и для поддержания существования вида...

Когда же происходят изменения, выбивающие организм из равновесия, которые не были предусмотрены «программой» инстинктов, реакции «регулятора первой ступени» оказываются недейственными, и начинается кризис. С одной стороны, резко повышается смертность неприспособленных организмов, и одновременно усиливается отбор, что дает преимущество определенным новым формам (мутантам)... С другой стороны, возникают исключительно благоприятные возможности для организмов, наделенных «регулятором второго типа», т. е. мозгом, который в зависимости от требований среды может изменять «программу действий» («самопрограммирование за счет обучения»)... Организм уже «на собственный страх и риск», не опираясь на готовую программу действия, либо приспосабливает себя к изменившейся среде (мышь учится находить выход из лабиринта), либо среду приспосабливает к себе (человек создает цивилизацию). Существует, разумеется, и третья возможность — «проигрыш»: создав ошибочную модель ситуации, организм не достигает нужного результата и гибнет».

Сопоставив эти высказывания С. Лема с анализом, проведенным нами в данном разделе, нетрудно заметить, что С. Лем описывает по существу тот механизм природы, который приводит к установлению оптимального соотношения случайных и детерминированных связей, выражаемого введенным параметром G_{opt}. Подчеркивая различия между детерминированными и стохастичными реакциями организмов, Лем говорит: «Организмы первого типа «все знают заранее». Организмы второго типа должны еще обучаться правильному поведению. Преимущества, которые дает первый тип «конструкции» организмов, оплачиваются их узкой специализацией, цена преимуществ организмов второго типа — риск».

Научный творческий поиск также всегда сопряжен с риском и неудачами (ошибочными пробами), которые не только в чисто финансовом, но и в более широком смысле оплачиваются весьма дорого (потеря времени, авторитета, веры в собственные силы и т. п.).

В процессе биологической эволюции все особи, отнесенные к категории «неудачных» проб, обречены на смерть. А поскольку, согласно терминологии Д. Т. Кэмпбелла, поиск методом проб и ошибок осуществляется путем «слепых вариаций», очевидно, что в общем количестве проб удачные составляют лишь ничтожно малую часть. Существующие в природе гармония и совершенство достигаются неумолимо суровым отбором и оплачиваются ценой массы смертей. Добиваясь в процессе эволюции оптимального соотношения случайных и детерминированных связей G_{opt}, природа сохраняет для организмов в форме мутаций ту долю риска, которая необходима для приспособления к изменчивым условиям внешней среды. Планируя технический прогресс и научный поиск, люди также стремятся найти G_{opt}, обеспечивающий поступательное движение науки и техники при необходимом минимуме рискованных проб.

Рассмотренный выше метод анализа соотношения случайных и детерминированных связей с помощью функции $\sum_i p_i \log p_i$ носит дедуктивный характер: вскрываются закономерности перехода от случайных связей к детерминации, общие для всех накапливающих информацию систем. Уяснив эти закономерности, мы проверяли их на частных примерах: на развитии инстинктов, письменной речи, общественных структур и т. п. К тем же выводам в принципе можно прийти и с помощью индуктивного метода: исследовать соотношения случайных и детерминированных связей в конкретных системах, а затем попытаться путем сопоставления выявить общие закономерности, не зависящие от конкретной природы систем. Проводимый системный анализ самых разнообразных объектов дает для таких обобщений обширный материал.

Например, в результате анализа ритмов поэтических произведений Ю. М. Лотман приходит к заключению, что ритмы стихотворных произведений представляют собой «игру упорядоченностей и их нарушений». «Поэтическая структура — прекрасная школа диалектики»,— замечает далее автор. Таким образом, на оснований исследований конкретных поэтических текстов Ю. М. Лотман делает тот же вывод, который был нами получен дедуктивным путем на основании описываемых функций

$$\sum_i p_i \log p_i ,$$

общих свойств эволюционирующих систем.

К этому следует лишь добавить, что отмеченная Ю. М. Лотманом «игра» безусловно осуществляется в пределах дозволенных правил, не допускающих разрушения гармонической структуры стиха. Эти пределы допустимой «игры упорядоченности и их нарушений» могут быть также выражены с помощью оптимального значения коэффициента стохастичности $G_{opt} = H_o/H_п$. При этом оптимальная величина H_o обеспечивает необходимые интонации и выразительность стихотворного текста, а оптимальная величина $H_п$ сохраняет музыкальную целостность стиха. Попытка количественных оценок соотношения стохастичности и упорядоченности музыкальных произведений была предпринята В. Детловсом. Установив фактические значения вероятностей чередований семи музыкальных аккордов, В. Детловс определил с помощью функции

$$\sum_i p_i \log p_i$$

значение величин, обозначенных нами как $H_{\max}$ и H_0. Вычисленный нами по данным В. Детловса коэффициент стохастичности чередований парных аккордов составляет G=2,4. Полученная величина близка к коэффициенту стохастичности чередований парных букв английского текста, вычисленному по данным К. Шеннона и составляющему G=2,3. Следующий пример, подтверждающий реальность существования оптимального коэффициента стохастичности G_{opt}, может показаться несколько неожиданным; тем не менее он имеет самое непосредственное отношение к процессам взаимодействия случайных и детерминированных связей. Речь идет о проведенных американскими социологами исследованиях процессов становления моды. После статистической обработки результатов опроса большого количества женщин было установлено, что $^3/_4$ опрошенных стремятся не отставать от моды с целью не отличаться от окружающих, а $^1/_4$ расценивают моду как средство, помогающее им выделиться из среды. Вполне очевидно, что именно эта $^1/_4$ часть выполняет творческую функцию «законодателей моды» и является в этой области источником непредсказуемой информации, оцениваемой введенным нами параметром H_o. Интересно отметить, что соответствующий этому процессу коэффициент стохастичности G = 1/3 весьма близок к выявленному нами оптимальному коэффициенту стохастичности языка. Совпадение свойств языка и моды не следует считать только забавным курьезом. С точки зрения системного подхода становление моды — это один из типичных эволюционных процессов, обладающий настолько быстрыми темпами, что в отличие от процессов биологической эволюции его результаты могут проявляться наглядно от сезона к сезону из года в год. Сделанный нами вывод о роли «мутаций» в процессе развития языка полностью совпадает с воззрениями пражской структурно-лингвистической школы, согласно которым структура языка как системы делится на основные участки («центры»), отличающиеся высокой степенью регулярности и упорядоченности, и «периферийные» зоны, в которых наблюдаются нарушения системности и проявляется динамика языка. Идея связи эволюции языка с нерегулярностями в его структуре (т. е. со стохастическими мутациями) получила развитие в работах представителя упомянутой школы — Р. Якобсона. Последний утверждает, что нерегулярности служат источником движения языка, в процессе которого «язык стремится ликвидировать состояние лабильности и перейти к новому гомеостэтическому состоянию». В соответствии с рассмотренными нами общими соотношениями состоянию гомеостазиса языка опять-таки соответствует оптимальный коэффициент стохастичности $G_{\text{opt}}=H_0/H_{\text{п}}$. Следующим примером, иллюстрирующим стремление эволюционирующих систем к полной детерминации, может служить подробно рассмотренный М. Ичасом генетический код. Исследования свойств наследственного кода показали, что сложившийся в результате эволюции «генетический словарь» обладает весьма ограниченными возможностями изменений. Ограничения эти обусловлены жесткой детерминированной связью между каждым из двадцати составленных из элементов генетического кода «слов» («кодонами») и включением в процесс синтеза белка одной из 20 аминокислот. Изменение любой из этих 20 связей порождает несовместимость комбинаций аминокислот в синтезируемом белке, что приводит к летальному исходу. В результате подобный мутант исключается из процесса дальнейшей эволюции, и полученное таким образом новое «слово» не может быть наследственно закреплено. Исходя из этого, М. Ичас приходит к выводу, что сам «генетический словарь» в настоящее время полностью детерминирован и неизменен, а мутации генов осуществляются теперь только путем изменения последовательности «слов». По его мнению, эволюция генетических кодов была возможна лишь на начальных этапах развития жизни, когда «вся биохимическая структура организмов была значительно менее специфичной» (в частности функционирование генома могло осуществляться без участия ряда ферментов), и соотношения между аминокислотами и кодонами были менее детерминированы («более лабильны»). В принятых нами обозначениях указанным условиям соответствует большое значение G за счет большого H_o и малого $H_{\text{п}}$. В ходе последующей эволюции «язык наследственности» достиг полной детерминированности на уровне «букв» (т. е. четырех нуклеотидов) и на уровне «слов» (т. е. выраженных генетическим кодом «наименований» 20 аминокислот). В результате на данном этапе развития земной биосферы эволюция осуществляется начиная с уровня «фраз» (путем изменения последовательности аминокислот). Вывод, сделанный М. Ичасом в результате анализа конкретных свойств генетического кода, полностью подтверждает справедливость общих соотношений, полученных нами на основе исследований «динамических» свойств функции

$$\sum_i p_i \log p_i,$$

проведенных в настоящем разделе. В самом деле, по мере увеличения количества «правил» возрастает детерминированность связей (в данном случае — детерминированность связей между кодонами и компонентами синтезируемых белков). К совершенно аналогичному выводу приходит В. В. Иванов, анализируя устную речь. На нижнем структурном уровне устная речь детерминирована определенным строением речевых органов, приспособленных для произнесения ограниченного числа фонем. Сравнительный анализ языков показывает, что для описания всех фонем, употребляемых во всех языках мира, потребуется набор признаков числом не более 20. Интересно отметить, что число фонем в человеческой речи (от 10 до 80 в зависимости от языка) имеет тот же порядок, что и число смысловых сигналов приматов (около 40). На основании этих данных В. В. Иванов делает естественный вывод о том, что «для развития человеческого языка... существенным было не увеличение числа самих первичных сигналов, а введение в систему иерархии разных уровней (фонем, морфем, слов, предложений), т. е. комбинаций первоначальных сигналов в последовательности и способ классификации по признакам...». Другими словами, анализ конкретных свойств разговорного языка приводит В. В. Иванова к выводу о том, что, достигнув определенной степени детерминации на уровне фонем, данная система начинает эволюционировать уже на более высоких структурных уровнях, т. е. ведет себя в полном соответствии с общими закономерностями эволюционных процессов, выявленными нами с помощью параметра G и изображенными наглядно на рис. 16.3.

Уменьшение потенциальных возможностей формообразований (изменений структуры) по мере роста $H_п$ и соответствующего уменьшения $H_o = H_{\max} — H_п$ (т. е. при малых значениях G) является общей закономерностью всех процессов самоорганизации, поэтому данную закономерность можно проиллюстрировать не только рассмотренными в предыдущем пункте языковыми примерами, но и сопоставлениями свойств конкретных биологических объектов. Так, например, рядом авторов отмечалась значительная изменчивость структуры микроорганизмов. «Даже незначительное изменение температуры среды или среды питания,— отмечает М. И. Сетров,— ведет к резким преобразованиям структуры, проявляющимся в изменении внешней формы клетки, ее размера, изменении ферментативных систем, самого принципа питания и дыхания. Вместе с тем оказывается, что для столь изменчивых в структурном отношении существ характерна высокая устойчивость физиологических функций. Противоположная картина наблюдается у организмов, стоящих на более высоких ступенях развития: ***здесь приспособление***

осуществляется путем изменения функции при сохранении той же структуры. Так, наиболее устойчивыми структурами высших организмов являются мышечная и нервная ткани. Но именно эти структуры и обладают, как писал Л. А. Орбели, «беспредельной потенцией к смене функций и их совершенствованию. Рост лабильности функций при сохранении прежней структуры можно проследить на примере роста в филогенезе мультифункциональности (по А. Н. Северцову) конечностей животных».

Интересно сопоставить свойства простых и сложных живых организмов со свойствами примитивных и совершенных языков. Примитивный язык — это язык, который еще не исчерпал «резервов мутаций», не накопил достаточного количества ограничивающих правил, что характеризуется малым $H_п$, большим $H_o = H_{max} — H_п$ и большим $G=H_0/H_п$. Вполне очевидно, что подобный язык приспособлен для выражения только самых простейших понятий, а усложнение понятий неизбежно сопровождается изменениями структуры языка (введением новых словообразований и грамматических структур). Этот процесс аналогичен отмеченному изменению структуры микроорганизмов, происходящих под действием окружающей среды.

Сложные организмы можно сравнить с высокоразвитыми языками, которым для выражения даже самых сложных понятий нет необходимости перестраивать свою структуру, так как они и при стабильной структуре обладают достаточным арсеналом выразительных средств. Это и есть не что иное, как «лабильность функций» совершенного языка и его «мультифункциональность», аналогичные тем же качествам, отмеченным А. Н. Северцовым и Л. А. Орбели для наиболее сложных и совершенных биологических структур.

Таким образом, рассмотренные результаты анализа конкретных систем в совокупности с приведенными ранее примерами (эволюционное образование детерминированных планетных систем, детерминированная структура клеток с переходом развития на уровень межклеточных связей, узкая специализация внутренних органов сложных организмов и др.) можно расценивать как эмпирическое подтверждение общих закономерностей эволюционных процессов, выявленных нами путем анализа динамических свойств функциии

$$\sum_i p_i \log p_i$$

и оцениваемых с помощью коэффициента *G*.

16.5. Две точки зрения на связь информации и энтропии

Вопрос о том, каким образом накопленная человечеством научная информация влияет на заложенные в фундамент науки идеи, не является прерогативой одного лишь данного раздела. В предыдущих разделах мы уже говорили о том, как рожденные в Древней Греции представления о роли случайностей впоследствии трансформировались в вероятностные трактовки взаимосвязи явлений. Мы проследили за тем, как наивная теория Эмпедокла, согласно которой случайные комбинации членов и органов подвергаются сортировке противоборствующими силами Любви и Вражды, со временем превратилась в апробированную практическим опытом теорию эволюции, в которой роль Любви и Вражды выполняют статистические взаимодействия организмов, популяций и видов с окружающей средой и их естественный отбор.

Мы убедились в том, что по мере развития науки накапливалось все большее количество фактов, подтверждающих, что понятия «случайность» и «вероятность» включают в себя не только субъективные факторы, связанные с неполнотой представлений о тех или иных явлениях, но и объективную стохастичность явлений. Мы отметили также, что вслед за ***понятием «вероятности» рожденная вероятностная мера информации проделала аналогичный путь, превратившись из субъективной меры неведения в объективную меру упорядоченности движения элементов самых разнообразных систем.***

Все это демонстрирует проявление одной общей тенденции — тенденции к непрекращающейся эволюции идей, обусловленной накоплением научной информации. В том же аспекте, прослеживая, с одной стороны, преемственность научных идей, а с другой — диалектическую противоречивость концепций, возникающих в ходе развития новых теорий, мы коснемся теперь вопроса об эволюции взглядов на природу и связь информации и энтропии.

Как отмечалось выше, первым качественным скачком в процессе становления понятия энтропии был переход от абстрактной «функции состояния» Клаузиуса к статистической интерпретации, которую дал этому понятию Больцман. Использование формулы Больцмана для оценки количества информации тех или иных сообщений сразу выдвинуло перед наукой новый вопрос: является ли такое использование функции $\sum_i p_i \log p_i$ чисто формальным математическим приемом или между количеством информации и физической энтропией существует некая глубокая связь? Точки зрения разделились. Многие ученые с полной определенностью высказались за то, что между физической и информационной энтропией следует признать лишь формальную связь. Противоположный взгляд, обоснованный Л. Бриллюэном, получил дальнейшее развитие в целом ряде трудов. Впоследствии Н. И. Жуков определил данную точку зрения как «расширительное толкование» понятия информации. Исходя из этого определения, утверждение о существовании только формальной связи между понятиями энтропии в физике и в теории информации правомерно назвать «ограничительным толкованием» информации. Согласно этой концепции, ***информация* — это особое свойство живой материи и искусственно созданных человеком технических средств.**

Начало «ограничительного толкования» положено не кем иным, как К. Шенноном, который в своих работах неоднократно подчеркивал, что теория информации преследует цели технической реализации оптимальных каналов передачи информации и вовсе не претендует на глобальный или тем более космогонический смысл. Стремление расширить полномочия теории информации, не ограничивая их рамками чисто технических приложений, К. Шеннон оценивал так: «Теория информации, как модный опьяняющий напиток, кружит голову всем вокруг... За последние несколько лет теория информации превратилась в своего рода бандвагон от науки...».

Однако вопреки воле и желанию главного создателя теории информации предложенная им мера количества информации помимо теории связи стала находить применение в биологии, физике, психологии, лингвистике и в ряде других областей. «Как это часто бывает в развитии пауки,— замечает по этому поводу Ю. В. Сачков,— решение, казалось бы, частной задачи приводит к разработке таких новых понятий и представлений, которые выходят далеко за рамки интересов исходной задачи и зачастую приобретают общетеоретическое значение. Так случилось и с понятием информации».

В результате развития теории информации и ее приложений основоположник этой теории Клод Шеннон оказался в роли персонажа известной сказки Шехерезады, который откупорил найденную им бутылку, не подозревая, какой огромный в ней скрывается Джинн.

Однако, как и в сказке, «Джинн информации» столь же могуч, сколь и миролюбив, а потому ждать от него следует только добра. И все-таки Клод Шеннон советует не слишком уж уповать на всемогущество Джинна, ибо «здание нашего несколько искусственно созданного благополучия слишком легко может рухнуть, как только в один прекрасный день окажется, что при помощи нескольких магических слов, таких, как «информация», «энтропия», «избыточность», нельзя решить всех нерешенных проблем...».

По поводу «всех нерешенных проблем» спорить, конечно, бессмысленно: было бы странно и удивительно, если бы все проблемы вдруг разрешились с помощью трех «магических» слов. Но среди «наболевших» проблем науки есть немало таких, к которым теория информации открывает новый подход. И дело тут не в «магической силе» слов «информация», «энтропия», «избыточность», а в глубине и емкости тех понятий, которые выражают эти слова. Рассмотренный нами выше ***процесс развития мира, трактуемый как процесс накопления информации***, сопровождающийся диалектическим переходом случайных связей в детерминированные — разве это не пример нового подхода к решению немаловажной проблемы?

Трудно предвидеть, какие вопросы удастся решить с помощью идеи информации в будущем, но уже сегодня становится очевидным, что информационная оценка явлений — это новый обобщающий взгляд на окружающий мир. Могучий Джинн выпущен из бутылки, и никакими заклинаниями спрятать его обратно уже нельзя.

В свете сказанного очень важным становится разрешение противоречий во взглядах на сущность понятия «информация» и на природу информационных процессов. Как справедливо отмечает Н. И. Жуков, «неоднозначность понимания этой важнейшей категории кибернетики наносит известный ущерб науке и практике», а «уяснению содержания и определение понятия — цель любого научного исследования в философии».

Можно провести прямую аналогию между современным этапом развития науки и техники и этапом, пережитым наукой и техникой около 200 лет назад. Если первая техническая революция, связанная с изобретением паровой машины, побудила науку установить единую сущность различных видов энергии, то вторая техническая революция, связанная с созданием электронно-вычислительных машин, поставила на повестку дня вопрос об углублении представлений о сущности, единстве и противоположности понятий информации и энтропии.

Вполне очевидно, что решение вопроса о справедливости одной из двух рассматриваемых нами концепций (т. е. «расширительного» или «ограничительного» толкования понятий «информации» и «энтропии») в значительной мере предопределяет путь дальнейшего развития исследований энтропийно-информационных соотношений различных процессов и ***методологию применения аппарата теории информации и статистической физики в исследованиях различных материальных систем.***

Попытаемся проанализировать шаг за шагом основные логические звенья «ограничительной» концепции информации на примере книги Н. И. Жукова, в которой делается попытка последовательно и полно обосновать указанную концепцию.

Исходным моментом проводимого Н. И. Жуковым анализа является утверждение о том, что информация не существует вне и помимо сложных саморегулирующихся систем. Автор книги предлагает определение информации как ««сведений» об окружающем мире, используемых организмом (или машиной) в своей деятельности». Развивая данную точку зрения, Н. И. Жуков приходит к выводу, что «информационные процессы возникли вместе с жизнью на Земле. ...Появление первого живого организма означало возникновение первой самоуправляющейся системы на Земле и используемой ею для управления информации». Для полной определенности остается лишь указать, в какой именно день творения был создан этот самый «первый» живой организм.

Противоположная точка зрения заключается в том, что между «живым» и «неживым», между «первым» и тем, который существовал до «первого», развивались какие-то переходные формы, подобные исследуемым современной наукой вирусам, способным в одних условиях проявлять все признаки жизни, а в других — вести себя как обычный кристалл. А поскольку существовали и существуют переходные формы от «неживого» к «живому», значит, должны существовать и переходные формы информации между информацией, участвующей, скажем, в процессах кристаллизации, и той информацией, которую использует в своей деятельности живой организм. При этом естественно считать, что, несмотря на многообразие переходных форм, информация сохраняет свою единую сущность. **<u>Передача информации</u> — это передача рассмотренных нами «моделей движения» (моделей колебаний атомов кристаллической решетки, передаваемых в процессе кристаллизации от кристалла к кристаллу, моделей развития организма, закодированных в цепочках ДНК, и т. п.). <u>Увеличение количества информации</u> в той или иной системе соответствует увеличению упорядоченности (детерминированности) статистического движения ее элементов.**

По нашему мнению, только такое толкование понятия информации — справедливое как для «живых», так и для «неживых» материальных объектов (хотя оно и отвергается Н. И. Жуковым и его сторонниками как «расширительное») — позволяет проследить связанный с накоплением информации процесс эволюционного усложнения материальных систем. Этот процесс начался на Земле задолго до возникновения первых живых организмов. Он проявлялся в форме передачи информации от молекулы к молекуле или от кристалла к кристаллу, приводя к накоплению информации упорядоченного движения атомов и молекул, образующих усложняющиеся в ходе развития неорганические, а затем органические вещества (кстати, заметим, что только благодаря способности к сохранению информации, присущей кристаллам, кристаллические вещества могут использоваться в качестве основных элементов электронных машин).

В соответствии с законами диалектики процесс количественного накопления информации сопровождался качественными скачками, а возникновение жизни как раз и явилось результатом одного из подобных скачков. При этом, конечно, возникли и качественно новые виды информации и новые способы ее передачи и накопления. Но **сущность информации (негэнтропии) как меры упорядоченности, меры отклонения от состояния равновесного хаоса сохраняется неизменной и для «живых», и для «неживых» материальных систем.**

Вопреки утверждениям сторонников «ограничительного толкования», усматривающих использующие информацию процессы авторегулирования только в деятельности живых организмов и созданных техникой автоматов, подобные процессы происходят и в целом ряде явлений неорганической природы (разумеется, в самой простейшей форме). В этих процессах можно обнаружить и информацию (негэнтропию), и прямую и обратную связь. В частности в процессе регулирования уровня воды в озере, который Н. И. Жуков упоминает как пример регулирования без участия обратной связи, в действительности функционирует обратная связь. Регулировка уровня данного озера — результат общего процесса авторегулирования, известного под названием круговорота воды в природе. Постоянство уровня в морях и озерах и непрерывность течения воды в реках обусловлены как пополнением водных бассейнов в виде осадков (связь прямая), так и испарением (обратная связь). Следуя дальше, можно найти и источник негэнтропии, определяющей упорядоченность (направленность) этого вида движения. Этим источником является нарушение термодинамического равновесия между Землей и Солнцем, обусловленное разностью температур. Кстати сказать, тому же источнику негэнтропии мы обязаны возникновением жизни, разума, электронных машин и других упорядоченных систем и процессов, образующих звенья единой цепи развития, происходящего на Земле.

Толкование энтропийно-информационных соотношений как диалектического единства двух противоположных сторон движения (стремления к хаотичности и упорядоченности) помогает связать все этапы развития материального мира в единую цепь. Сторонники «ограничительного толкования» информации по существу не выдвигают против такой точки зрения каких-либо аргументов.

Точка зрения, согласно которой информация есть мера материальных систем, отвергается, например, Н. И. Жуковым потому, что «подобное расширительное толкование в итоге может привести, например, к ошибочной трактовке кибернетики, как науки об отражении, и возведению информации в ранг философской категории». Эта аргументация не представляется, однако, убедительной.

Кибернетика, как известно, представляет новое направление научной и технической мысли, основывающееся на синтезе «старых» и «новых» математических дисциплин. Теория же отражения входит в философию диалектического материализма. Поэтому можно считать, что кибернетика соотносится с теорией отражения так же, как физика и математика соотносятся с философией. Результаты кибернетики и теории информации служат развитию положений материалистической теории отражения.

Исходя из общего понимания информации как статистической меры упорядоченности (напрвленности) движения, роль и место понятия информации в современном научном мировоззрении можно определить следующим образом. «Ранг» информации совпадает с «рангом» понятия энергии, поскольку и энергия, и информация служат мерами, характеризующими различные формы движения с двух разных сторон. Разумеется, ни энергия, ни информация не могут претендовать на «ранг» философских категорий, поскольку это естественнонаучные понятия, позволяющие исследовать различные формы движения, а из всех упомянутых выше понятий только последнее (движение) может быть отнесено к разряду философских категорий.

В стремлении ограничить «сферу влияния» информации Н. И. Жуков заходит так далеко, что вводит понятие «биологической энтропии», противопоставляя ее общепринятому понятию термодинамической энтропии. Между тем, на наш взгляд, наиболее правильный подход к изучению объектов живой природы должен базироваться на признании того факта, что живой организм безусловно подчиняется всем физическим законам, включая и законы термодинамики. Это, разумеется, ни в коей мере не означает, что на базе физических законов в живой природе не возникают и более сложные биологические законы (например, законы синтеза белков или обмена веществ, законы нейро-гуморального регулирования физиологических процессов и т. п.). Но речь сейчас идет не об этих специфических биологических законах, а о приложимости общих законов термодинамики к исследованию биологических объектов.

При исследованиях биологических объектов речь может идти не о «биологической энтропии», а о *термодинамической энтропии биологических систем.* Для установления связи между процессом накопления информации живым организмом и его способностью противостоять общей тенденции к увеличению энтропии как раз и необходимо введение единых и одновременно противоположных статистических мер упорядоченности и хаотичности движения, каковыми являются меры количества информации и энтропии.

Анализ диалектического единства понятий информации и энтропии как одной из форм борьбы противоположностей в движении материального мира (т. е. стремления к увеличению хаотичности и к упорядоченности движения), так же как установление общности энтропийно-информационных соотношений в процессах развития любых материальных систем, является актуальной задачей современного естествознания. Данное утверждение, конечно, отнюдь не исключает отстаиваемую сторонниками «ограничительного толкования» необходимость разграничения форм информации и разработки специальных методов для исследования этих форм. На необходимость классификации видов информации (информация «связанная» и «свободная», «живая» и «мертвая» и т. п.) указал уже один из первых сторонников «расширительного толкования» информации — Л. Бриллюэн. Несостоятельна, на наш взгляд, и попытка опровержения негэнтропийного принципа информации, установленного Л. Бриллюэном, на том лишь основании, что Бриллюэн опирается на теорию Шеннона, которая при количественных оценках информации игнорирует ее ценность и смысл. «Некорректность выражения «негэнтропийный принцип информации»,— пишет Н. И. Жуков,— очевидна, если учесть, что в техничсской кибернетике термин «энтропия» часто употребляется в смысле математической оценки информации (в выражении «энтропия информации» у Шеннона, например). Этот математический аппарат никак не учитывает значимости, ценности информации, что признает даже тот, кто вышеназванный принцип склонен считать главным в теории информации». «...Назвать информацию негэнтропийной — значит игнорировать тот факт, что формула количества информации, предложенная Шенноном: $I = -\sum_i p_i \log p_i$. принципиально не подходит для определения семантической информации».

Согласно приведенным рассуждениям, негэнтропийный принцип информации не может быть признан справедливым по той причине, что оценка количества информации производится независимо от ее смыслового значения. На самом деле все обстоит как раз наоборот. Всякая семантическая оценка информации в том или ином аспекте в конечном счете субъективна. **Главным достижением теории информации является именно то, что она сумела абстрагироваться от семантической стороны информации и благодаря этому нашла *объективную количественную меру*.** Без этой абстракции наука никогда не смогла бы перешагнуть за пределы традиционных рамок узкобытового понятия «информация», интерпретируемого как передача и получение смысловых новостей. **Именно благодаря абстрагированию от конкретных свойств информации, ее ценности, смысла, науке удалось сделать первый и решающий шаг в исследовании закономерностей энтропийно-информационных процессов, подобно тому как абстрагирование от конкретных свойств механического движения (неизбежности трения) в свое время позволило Галилею сделать в изучении общих законов механики столь же решительный шаг.**

Вопросы о смысле и ценности информации, о ее материальных носителях, о способах обработки и т. д. — все это хотя и очень важные, но все же частные вопросы, возникающие при анализе конкретных процессов. А ***оценка количества информации и энтропии — это общий метод сравнительной оценки степени организованности, направленности, упорядоченности любых материальных систем***.

(Точно так же закон сохранения энергии — *общий* закон, а вопросы о формах энергии, ее передаче, преобразованиях и т. п.— частные вопросы исследования конкретных систем.)

Семантическая оценка информации — это один из видов селекции. Можно сортировать информацию по ее смыслу, ценности, содержанию. (Так, в частности, поступает руководитель любой организации, когда он по смысловому значению информации направляет полученные письма на исполнение Иванову или Петрову, а утратившие актуальность адресует в архив.) Но сортировать информацию можно и по свойствам ее материальных носителей: запах, несущий собаке огромное количество информации, не воспримется человеком; приемник, настроенный на волну определенной станции, не воспримет информацию, передаваемую в другом диапазоне частот. Все это — примеры селекции. И хотя смысловая селекция несравненно сложнее и важнее, чем другие формы селекции (ведь «фильтром» здесь служит сознание!), по отношению к общему процессу развития, отбора и накопления информации все формы селекции функционально стоят в одном и том же ряду. Это не значит, конечно, что их изучение должно непременно проводиться с помощью единого универсального инструмента. Напротив, к исследованиям различных форм селекции необходим разный подход. Так, например, частотная селекция информации, осуществляемая в многоканальных системах связи, основана на исследовании частотных и резонансных характеристик каналов. Для исследования упомянутой выше селекции информации с помощью органов обоняния необходимо прежде всего изучить физическую природу запахов и т. п.

Вопросы селекции информации по ее смыслу и ценности относятся к числу наиболее сложных, поэтому вполне естественно, что в этой области пока сделаны лишь первые шаги. Предложены не вполне совершенные логические модели, позволяющие производить количественные оценки информации с учетом ее содержания («семантическая теория информации» или ценности («прагматическая теория информации»).

Предпринятые многими авторами попытки отрицания общенаучного «ранга» понятия «информации» основываются иногда и на том, что предложенный К. Шенноном для оценки количества информации аппарат недостаточно универсален. Если дело касается вопроса о конкретных исследованиях различных видов информации и различных форм информационных процессов, то тут, безусловно, приходится согласиться с утверждением об ограниченной применимости формулы

$$I = -\sum_i p_i \log p_i.$$

Так, например, в указанном виде в данной формуле учитываются вероятности появления разных сигналов в отвлечении от их взаимосвязи. Чтобы учесть корреляцию, выражение количества информации приходится несколько усложнить. Еще сложнее производить количественные оценки информации с учетом ценности и смысла. Например, А. Н. Колмогоров считает, что в этом случае величина *I* должна быть не скалярной, а векторной. Дальнейшее развитие кибернетики, теории информации, статистической физики, биологии, теории случайных процессов позволило найти новые методы исследований различных видов информационных процессов. Однако с точки зрения современного научного мировоззрения **главная заслуга теории информации заключается именно в том, что она позволила установить во всем многообразии видов и форм энтропийно-информационных процессов их единую сущность и взаимосвязь**.

Четкое обоснование общенаучного смысла понятия «информация» дает Ю. В. Сачков. Общетеоретическое значение этого понятия Ю. В. Сачков объясняет тем, что «оно служит для характеристики и исследования весьма широкого класса множеств безотносительно к их конкретной природе, будь то множества сообщений или материальных объектов и сущностей. В основе развития теории информации лежит анализ множеств, имеющих вероятностную структуру, и она выступает прежде всего как своеобразное «исчисление» вероятностных структур... Соответственно этому **информация обобщенно определяется через представление о разнообразии и неоднородности в строении и движении материи»**.

Попутно заметим, что упомянутая в данном высказывании Ю. В. Сачкова характеристика информации как меры неоднородности распределения энергии в пространстве и во времени полностью соответствует определению, данному Е.А.Седовым и подробно рассмотренному в данной работе. В самом деле, характеризуемое количеством объективно сохраняемой информации упорядоченное (направленное) движение возникает именно благодаря неоднородности распределения энергии в пространстве и во времени, и в этом смысле определения «информации», предложенные Е.А.Седовым и В. М. Глушковым, равноценны. Различие заключается лишь в том, что рассматриваемое Е.А.Седовым представление о количественной мере информации как мере упорядоченности (направленности) движения устанавливает непосредственную связь между информацией и материальным объектом, в то время как В. М. Глушков определяет информацию через вторичную характеристику движения (через энергию). Понимание информации как меры упорядоченности (направленности) движения представляется нам более точным, в частности, потому, что оно позволяет сравнить по степени общности понятия энергии (меры интенсивности движения) и информации (меры упорядоченности движения) и поставить их в один ряд.

Отрадно отметить, что тенденция расширительного толкования понятия информации в последние годы получает все более широкое признание в научной среде. По этому поводу В. М. Адров писал: «...Информацию мы будем понимать в несколько более широком смысле, нежели это делается в кибернетике. Предпосылки для этого, в частности, дает и сама теория информации, все основные построения которой создавались весьма независимо от идей кибернетики. И если основные результаты теории информации нашли самое широкое применение в кибернетике, то это означает лишь то, что кибернетические системы (в том числе и живые существа) не могут существовать без информационных процессов, но не наоборот.

Глубокий философский смысл понятия информации, который сейчас признается всеми, как раз и связан с тем, что понятие информации можно расширить за пределы кибернетики и связать с философскими категориями». Ту же точку зрения отстаивает и развивает в своей книге А. Д. Урсул.

16. 6. Диалектика случайных и детерминированных связей

Среди результатов и выводов, получаемых на основе кибернетических и информационных концепций, принципиальное значение имеет вывод о необходимости введения в программу моделирования процессов эвристического мышления с помощью электронно-вычислительных машин стохастического сигнала. Впрочем, идею о роли стохастических связей в процессе творческого мышления не следует относить к числу заслуг одной лишь кибернетики, поскольку эту мысль высказал еще в 1855 г. А. Бэн. Ему же, кстати, принадлежит и термин «метод проб и ошибок», получивший новое звучание и широкую популярность опять-таки благодаря использованию этого метода в программах электронных машин. Вслед за А. Бэном другие авторы также рассматривали творческое мышление как стохастичный процесс.

Таким образом, роль кибернетики в данном случае сводится к развитию ранее высказанных идей. Однако в общем аспекте гносеологии кибернетический подход весьма принципиален: то, что во времена А. Бэна и его ближайших последователей было всего лишь гипотезой, теперь превратилось в экспериментальный факт. Доказано, что без введения в программу ЭВМ стохастического сигнала нельзя осуществить эвристического, «творческого» решения задач. Естественно возникает желание использовать опыт моделирования на электронных машинах с целью уточнения представлений о реальных процессах мышления. Именно этим обстоятельством обусловлено опубликование многочисленных работ, посвященных рассмотрению процессов мышления в кибернетическом аспекте. В этих работах высказано немало интересных и заслуживающих внимания гипотез о природе мышления. Но вместе с тем в ряде случаев имеет место тенденция к упрощенным трактовкам такого сложного явления, как человеческая мысль. Именно с таких позиций пытается анализировать процесс эвристического мышления Д. Кэмпбелл.

Лейтмотивом работы Кэмпбелла является утверждение, что в основе творческого поиска лежат *слепые вариации,* не зависящие ни от прошлого опыта, ни от цели поиска и поставленных творческих задач.

По мнению Кэмпбелла, ни удачно найденное решение проблемы, ни открытие, ни изобретение—«ничего не говорит нам о превосходстве гения одного человека — просто так случилось, что он стоял на том месте, которое внезапно озарила молния». «Мы испытываем соблазн,— рассуждает Кэмпбелл, искать у этого человека наличие особого неуловимог таланта... В случае подлинно не поддающегося предвидению творческого акта наше «благоговение» и «удивление должно быть направлено вовне, к внешнему миру, который так себя обнаруживает, а не к тому, что предшествовало этому открытию. Точно так же, как мы не приписываем особого «предвидения» удачному мутанту по сравнению с неудачным, мы во многих открытиях не должны предполагать, что изумительные заключения имели столь же изумительное прошлое».

Надо признать, что сравнение с удачным мутантом выбрано неудачно. «Удачный мутант» — это тоже результат всей предшествующей эволюции, которая как раз и является «изумительным прошлым» биосферы Земли. Учет этого «изумительного прошлого» — необходимое условие анализа объективных свойств процесса эволюции биологических видов.

Простой расчет, проведенный М. Иденом, показывает, что путем «чисто случайного» перебора комбинаций невозможно создать не только сложный организм, но даже молекулу белка. Если принять, что среднее число единиц в полипептидной цепи равно 250, то число комбинаций равно числу «слов», написанных с помощью алфавита из 20 букв (т. е. 20 различных аминокислот), при условии, что каждое «слово» состоит из 250 букв. Число таких «слов» составляет 20^{250} или 10^{325}. Общее число когда-либо существовавших белковых молекул по приблизительным подсчетам Идена составляет около 10^{52}. Вероятность того, что одно случайное «слово» из 10^{325} «слов» окажется белковой молекулой, составляет ничтожную величину: $10^{52}/10^{325}=10^{-273}$. Иден приходит к выводу, что реализация подобной случайности даже в течение миллионов столетий практически столь же невероятна, как невероятно, чтобы ребенок, случайно переставляющий буквы, вдруг набрал из «Энеиды» Вергилия хотя бы первые 20 строк.

Однако с точки зрения Кэмпбелла нечто подобное происходит в процессе творческого мышления: при решении той или иной научной проблемы шансы на успех у исследователя, наделенного эрудицией и опытом, ничуть не выше, чем у того, кто пытается ту же проблему решать «с кондачка». (Если продолжить аналогию с мутантами, это равносильно утверждению о том, что лебедь мог с равной вероятностью произойти и от археоптериксов и от плоских червей.) Кэмпбелл полагает, что благодаря слепым вариациям «многие важные вклады в науку будут сделаны сравнительно неталантливыми и малоусердными людьми». Иная позиция, по мнению Кэмпбелла, «приводит к «обожествлению» творческого гения, которому мы приписываем способность непосредственного ясновидения, вместо барахтанья мыслей и блуждания по тупикам, что является, как мы это осознаем, прообразом наших процессов мышления».

Впрочем, сам Д. Кэмпбелл чувствует шаткость своих заключений. Не случайно он вспоминает, что еще в 1726 году Свифт создал пародию на подобное толкование творческого процесса, описав в «Путешествиях Гулливера» лапутянскую академию, ученые которой пытаются открыть великие истины, нанизывая на стержни случайные последовательности букв. Упоминается и современная модификация этого «метода», которому группой авторов было присвоено шуточное название «алгоритма Британского музея». Суть его заключается в том, чтобы заставить обезьян стучать на пишущих машинках. Расчеты показывают, что путем случайного перебора букв в течение миллиона лет группа обезьян наряду со множеством бессмысленных вариаций создаст и осмысленные страницы всех книг, которые хранятся в библиотеке Британского музея.

Так в чем же все-таки состоит отличие талантливого ученого от обезьяны? Пытаясь ответить на этот вопрос, Кэмпбелл приходит к выводу, что разница эта проявляется не на этапе поиска, а лишь на этапе отбора «слепых вариаций». «Различие между успехом и неуспехом лежит в условиях выбора встретившихся заново комбинаций, а не в различиях талантов при создании проб»,— утверждает Кэмпбелл. Сокращение времени поиска по сравнению с «творчеством» по «алгоритму Британского музея», по мнению Кэмпбелла, достигается тем, что отбор неудачных проб производится последовательно, шаг за шагом, «с последующим испытанием только тех вариаций, которые прошли отбор на каждой предыдущей стадии». Для экономии времени «в процессе создания нашей «универсальной библиотеки» мы прекращаем работу на любом томе, как только становится ясным, что он представляет собой тарабарщину».

Итак, процесс творческого поиска Кэмпбелл подразделяет на две стадии, из которых первая *абсолютно случайна* (поиск), а вторая — *детерминирована* (отбор). Согласно трактовке Кэмпбелла, такие факторы, как способности, эрудиция, опыт, проявляются лишь на второй стадии, в то время как первая стадия совершенно от них не зависит.

«Каковы же признаки, по которым можно различать мыслителей друг от друга, исходя из концепции модели проб и ошибок? — спрашивает Кэмпбелл.— Прежде всего, они могут отличаться друг от друга точностью и подробностью своих представлений об окружающем мире, о манипуляциях с его элементами, а также представлений о критериях выбора... Творческий мыслитель большого масштаба может удерживать в уме большое количество таких критериев, и поэтому увеличивается вероятность его успеха в решении проблемы, соприкасающейся с первоначальным главным направлением его попыток. Последняя область индивидуальных различий в способностях связана с умением накапливать и передавать встречающиеся решения».

Из своих рассуждений Кэмпбелл делает практический вывод: «Эти соображения наводят на мысль о целесообразности взаимно дополняющих друг друга комбинаций талантов в творческих объединениях, хотя, как известно, несдержанный, обуреваемый идеями человек и человек методического склада, склонный редактировать и протоколировать, являются плохо совместимыми коллегами».

Методологической основой взглядов Кэмпбелла служит постулат, утверждающий, что стохастический поиск и детерминированная направленность несовместимы. На деле **эти оба процесса могут *диалектически совмещаться в одном мозге, в одно и то же время* — реализовывать диалектически *противоречивый,* но в то же время *единый* процесс.**

Источником выводов, подобных заключениям Кэмпбелла, является не преодоленная полностью современной наукой метафизичность подхода, при котором «какая-нибудь вещь, какое-нибудь отношение, какой-нибудь процесс либо случайны, либо необходимы, но не могут быть и тем и другим». Между тем именно совмещение случайных и детерминированных связей может быть залогом и необходимым условием всякого творческого успеха. Тот факт, что такое совмещение в пределах единой системы в принципе возможно, подтверждает анализ статистических свойств языка, проделанный нами в настоящем разделе выше.

Отсутствием диалектичности в подходе к анализу процесса мышления грешит, к сожалению, не только Кэмпбелл. С тех же самых позиций анализирует процесс мышления У. Росс Эшби, утверждающий, что «в гении замечательно умение отсеивать возможности». Поэтому способность выбрать оптимальную возможность из N заданных Эшби считает чуть ли не абсолютным показателем развития ума. Так же как и Кэмпбелл, Эшби искусственно расчленяет процесс мышления на две части: «селективную» и «шумовую». На этой основе он высказывает гипотезу о принципиальной возможности создания «усилителя мыслительных способностей», который в представлении Эшби есть не что иное, как селектор генерируемых опять-таки в виде шума идей.

Существование подобных «неометафизических» концепций лишний раз подтверждает актуальность диалектического анализа взаимодействия случайных и детерминированных связей в процессах мышления, начиная с единичных творческих актов и кончая общим развитием естественнонаучной мысли. Такой анализ раскрывает, что, как и во всяком процессе диалектического развития, в процессе творческого мышления осуществляются переходы случайных связей в детерминированные, рассмотренные нами на примере эволюции языка.

Поскольку язык является средством выражения мыслей, нетрудно себе представить, что подобное противоречивое единство случайных и детерминированных связей присуще как процессу мышления, так и тем информационным связям между клетками коры мозга, которые обеспечивают этот процесс. Напротив, бессистемное, нецеленаправленное «блуждание по тупикам» и «барахтанье мыслей», отстаиваемые Кэмпбеллом, разумеется, не могут принести ничего, кроме бесплодной растраты времени и творческих сил.

Проводя аналогию между эвристическим мышлением и биологической эволюцией, Кэмпбелл, казалось бы, уже готов признать диалектическое единство их случайных и детерминированных свойств: «При развитии органического мира процесс мутационного изменения и процесс сохранения приобретенных свойств посредством консервативной наследственности находятся в некотором противоречии друг с другом, причем увеличение любого из этих факторов происходит за счет другого, а некоторая степень компромисса между ними является оптимальной». Казалось бы, этот вывод должен привести Кэмпбелла к анализу взаимодействия случайных и детерминированных связей. Однако этот автор идет по другому пути: «Мы могли бы ожидать, что очень ярко выраженный диапазон новаторства в мышлении и очень ярко выраженная механическая память могли бы привести даже к отрицательной взаимной корреляции». И далее следует уже приведенное нами высказывание о том, что единственный выход из этого противоречия заключается в подборе «взаимно дополняющих друг друга комбинаций талантов».

Таким образом, Кэмпбелл проходит мимо диалектического единства противоположностей («некоторой степени компромисса»), которое приводит к оптимальному соотношению случайных и детерминированных связей, характеризуемому введенным параметром G_{opt}. Между тем, так же как и в любом процессе самоорганизации, такой компромисс (а точнее — оптимум) существует и в процессе эвристического поиска. Практически он проявляется, во-первых, в правильном выборе некоторого количества (но не бесконечно большого!) вероятных направлений успешного поиска, во-вторых, в правильном определении возможного диапазона поиска ассоциаций между явлениями, понятиями и т. д. Чрезмерное (т. е. далекое от оптимального) сужение диапазона и направлений случайного поиска закрывает путь к неожиданным ассоциациям, или, как выражаются современные физики, «сумасшедшим идеям». Но, с другой стороны, успеха в творчестве достигает не тот, чья мысль «барахтается» и «блуждает» без всяких ограничений, а тот, кто сумел априорно определить правильные направления поиска и объективные критерии оценки того, что в каждом из выбранных направлений удается найти. Лишь после того, ***как четко сформулирована цель научного поиска и намечены возможные пути достижения поставленной цели, можно осуществлять в выбранных направлениях стохастичный поиск, не слишком сужая его границ.***

В умении совместить оптимальным образом противоречивые условия стохастичности и априорной детерминации и заключается истинный творческий талант. **Индивидуум, мыслительный процесс которого характеризуется соотношением $G \ll G_{opt}$, не способен в своем творчестве выйти за рамки общепринятых методов, представлений и взглядов. Индивидуум, у которого $G \gg G_{opt}$, обречен вечно терзаться обилием преимущественно бесплодных идей. Бесплодность его идей обусловлена тем, что в силу отсутствия необходимой степени детерминированности интеллекта (дисциплины мышления) он не способен отличить важные идеи от незначительных соображений и направить по нужному руслу свою «шарахающуюся» и «блуждающую» мысль. И лишь при условии $G \cong G_{opt}$ может плодотворно осуществляться продуктивный творческий процесс.**

Согласно преданию, падение яблока «случайно» натолкнуло Ньютона на открытие закона всемирного тяготения. Существует версия, согласно которой периодическая система химических элементов Менделееву «случайно» приснилась во сне. Майкл Фарадей «случайно» обнаружил, что для получения тока индукции необходимо либо изменять магнитное поле, либо перемещать магнит относительно проводника. Клод Шеннон «случайно» решил применить статистическую формулу физической энтропии для оценки энтропии и информации передаваемых сообщений, а потом оказалось, что формула эта пригодна и для оценки количества информации, сохраняемой и в тканях живых организмов, и в структуре физических тел. Во всех приведенных примерах «случайным» открытиям предшествовала гигантская целенаправленная работа творческой мысли.

Можно привести много примеров того, как важные открытия и изобретения рождались путем неожиданных сопоставлений несходных явлений из далеких научных или технических областей. Этим процесс мышления отличается от процесса биологической эволюции, в котором многообразие видов и видовых признаков обусловлено не широким диапазоном мутаций, т. е. не большим значением дисперсии признаков у одного поколения, а последовательным накоплением небольших отклонений в течение многих сменяющихся поколений. Именно благодаря этому свойству возникает тенденция постепенного «целенаправленного» развития органов. Именно в этом смысле следует понимать замечание К. Уоддипгтона: «Мы, разумеется, не должны считать, что глаз позвоночного животного, нога лошади или шея жирафа представляют собой в сколько-нибудь серьезном смысле результат случайного поиска».

Мысль не имеет подобных ограничений. Она может мгновенно «мутировать» ассоциации между любыми явлениями, и именно благодаря этому наше сознание способно за ограниченное время строить модель таких процессов, как эволюция биологических видов, реальная продолжительность которых составляет миллионы и миллиарды лет. Однако это вовсе не означает, что подобные ассоциации возникают «чисто случайно», что на них нечаянно «натыкается» «слепая», «блуждающая» мысль. Напротив, они всегда являются результатом целенаправленного поиска, сопоставления приобретенных предварительно глубоких знаний из нескольких не связанных между собой, но осознанных данным исследователем областей.

Именно эту целенаправленную, детерминированную составляющую творческих поисков ученого А. А. Ухтомский называл «доминантой». **Под доминантой научного мышления он понимал длительную установку научного поиска, своего рода «сверхзадачу» ученого, или, как выражался сам Ухтомский, его «маховое колесо».** При этом А. А. Ухтомский подчеркивал, что чрезмерное преобладание доминанты приводит к консерватизму научных взглядов, к неумению уяснить для себя доминанты своих коллег: «Человек является настоящею жертвою своих доминант везде, где отдается предубеждению, предвзятости; и еще хуже, когда он сам этого не замечает. **Чтобы не быть жертвою доминанты, надо быть ее командиром**. По возможности полная подотчетность своих доминант и стратегическое умение управлять ими — вот практически что нужно».

Отмеченному А. А. Ухтомским «стратегическому умению» управления доминантой соответствует определенное соотношение случайных и детерминированных связей мышления и соответствующий ему оптимальный коэффициент стохастичности G_{opt}. **Только благодаря наличию соответствующих G_{opt} вероятностных связей, выражающихся количественно величиной $H_п$, индивидуум сохраняет способность воспринимать неожиданную информацию даже в том случае, если она противоречит его доминанте, и пересматривать собственные сложившиеся представления с учетом чужих концепций, взглядов, идей.**

Другим примером сочетания случайных и детерминированных связей в процессах эвристического мышления является интуитивный творческий поиск. Общепризнано, что в успехе решения любой новой научной проблемы интуиции исследователя отводится далеко не последняя роль. А что представляет собой интуиция, как не совмещение случайных и детерминированных процессов, когда накопленный опыт подсказывает примерное направление поиска, а недостаток исходных данных восполняется удачными решениями, принимаемыми «наугад»?! В этой связи следует упомянуть предложенную некоторыми авторами одностороннюю трактовку понятия «интуиции» как случайного выбора решений, которая сводит это ценнейшее качество творческой личности к «автоматической интуиции», имитируемой генератором случайных сигналов, вводимых в эвристическую программу электронных машин.

Если бы каждый научный поиск был полностью стохастичным, это означало бы равенство вероятностей всех направлений поиска. Такой ситуации соответствует полное отсутствие предварительной информации, т. е. максимально энтропийный поисковый процесс. Так может «творить» обезьяна, для которой не существует смысла возникающих слов. Ребенок, овладевающий навыком чтения, творит и ищет уже направленно: на основе усвоенной разговорной речи он стремится составить из букв произносимые слоги и осмысленные слова. **Предварительный опыт ученого позволяет ему осуществлять частично направленный, а частично случайный (по методу проб и ошибок) поиск новых связей между явлениями, понятиями и т. д.** Даже в таких упрощенных моделях эвристического процесса, каковыми являются на сегодняшний день самообучающиеся программы для ЭВМ, успех обучения может быть достигнут только в том случае, если машина умеет «детерминировать» поиск, запоминая как удачные, так и неудачные пробы, а затем сортируя их в соответствии с определенными критериями таким образом, чтобы неудачные пробы в дальнейшем не повторять.

При отсутствии определенной детерминации творческого процесса «поиск» состоял бы в выборе первого попавшегося пути. К счастью, **в отличие от длящейся тысячелетиями «слепой» биологической эволюции, <u>творец</u> — художник, ученый и т. п. — <u>всегда ставит определенную цель.</u>** Это дает возможность использовать как индуктивный, так и дедуктивный подход, т. е. идти как от частного к общему, так и от общего к частному, сформулировать в общих чертах ту мысль, которой он посвятит следующую строку поэмы, а уж затем решать более частные вопросы относительно необходимых для выразительности и одновременно удовлетворяющих требованиям стихотворных размеров и рифмы отдельных слов.

Как было показано выше, наиболее ярким художественным образам («страна березового ситца» и т. п.) соответствует оптимальный коэффициент стохастичности G_{opt}. При оценке степени направленности (детерминированности) творческого поиска следует учитывать еще и то обстоятельство, что различные его направления всегда в той или иной степени взаимозависимы, ибо неудача в одном из направлений помогает найти правильный путь в другом. Это является дополнительным фактором, уменьшающим исходную неопределенность поиска, так как в соответствии с исследованными теорией информации свойствами функции $\sum_i p_i \log p_i$ неопределенность H уменьшается в случае зависимых значений p_i (в этом случае символы p_i следует понимать как *условную вероятность*). Априорное определение оптимальных соотношений случайных и детерминированных связей в процессе творческого поиска становится актуальной практической задачей в связи с назревшей необходимостью планирования научно-технического развития.

Ввиду отсутствия методов априорного определения оптимального соотношения случайных и детерминированных связей правильность выбора этого соотношения может быть проверена только путем практической реализации и координации научно-исследовательских работ. Поиск разрешения противоречия между стохастичностью и детерминированностью осуществляется и при решении технико-экономической проблемы: определения оптимальной степени унификации и стандартизации технических средств. Очевидно, что унификация приносит экономический выигрыш, однако одновременно она же тормозит внедрение новых решений и средств. Определение разумных пределов стандартизации и унификации — это еще один пример поисков оптимального значения коэффициента стохастичности G_{opt}.

Анализируя роль случайного поиска в процессе биологической эволюции, Уоддингтон повторяет ошибку Кэмпбелла, стремясь во что бы то ни стало расчленить на «чисто случайную» и «жестко детерминированную» части диалектически единый процесс. По его мнению, случайный поиск может осуществляться только на самом первичном уровне организации «генетического материала» и выражается в добавлении, исключении или изменении последовательности нуклеотидов в ДНК. «Следует ли из этого, что изменения, на которые действует естественный отбор, также случайны?» — спрашивает Уоддингтон. А в качестве ответа приводит следующий пример: «Форма гравия на дне реки определяется случайными процессами, т. е. возникает в результате случайного поиска. Однако из этого не следует, что случайный поиск играет сколько-нибудь существенную роль в строительстве моста из бетона, для приготовления которого был использован этот гравий. Факторы, учитываемые при сооружении моста, относятся, так сказать, к иному порядку сложности, чем те, которые определяют образование компонентов бетона. Надо спросить, не обусловлена ли приписываемая эволюции зависимость от случайного поиска сходным смешением различных порядков сложности?».

В процессе биологической эволюции «гравием» является упомянутый выше «генетический материал», для которого Уоддингтон признает право на случайные мутации отдельных генов. Однако «у высших организмов... изменения, оказавшиеся выгодными в эволюционном плане, зависят вообще не от случайных мутаций единичных генов. Преобладающее большинство случайных генных мутаций, имеющих достаточно выраженный эффект, чтобы их можно было заметить по отдельности, оказываются вредными и элиминируются естественным отбором,... Включение мутантной формы какого-либо из главных генов в эволюционную последовательность влечет за собой одновременный отбор большой связанной с ним группы «генов-модификаторов». Однако в общем эволюция высших организмов зависит от отбора признаков, на которые более или менее одинаково интенсивно влияет большое число генов. Такие признаки можно сравнить скорее с бетонными блоками, а не с отдельными камешками гравия в бетоне».

Эти слова не вызывали бы никаких возражений, если бы Уоддингтон признал право случайного поиска и для строителей указанного «моста». Но «мост Уоддингтона» почему-то должен быть непременно построен по стандартному *детерминированному* проекту, из которого случайный эвристический поиск должен быть полностью *исключен.* «Возвращаясь к аналогии с бетонным блоком, можно сказать, что роль случайных процессов заключается в том, чтобы «создать» гравий; затем мы можем из этого гравия приготовить бетон, а этому бетону придать форму какого-либо объекта, приспособленного к окружающей среде. Эта форма не будет воспроизводить объект во всех деталях, если гравий был слишком крупный; если же гравий, напротив, был слишком мелкий, то форма будет недостаточно прочной. Самое главное — это чтобы смесь имела оптимальный состав. В образовании гравия участвуют лишь случайные процессы, но от этого еще очень далеко до вывода, что формы, образующиеся в соответствии с тем или иным объектом, т. е. фенотипы, в разных условиях среды также возникают путем случайного поиска».

Мы видим здесь тот же подход, что и у Кэмпбелла: разрыв единого процесса, выделение из него «чисто случайной» и «жестко детерминированной» частей. Но выводы прямо противоположны: Кэмпбелл преувеличивает роль случайного поиска в процессах эвристического мышления, Уоддингтон умаляет роль случайных мутаций и переоценивает роль детерминированных связей в биологических эволюционных процессах.

Тенденция преувеличения роли одной из сторон явлений (детерминированности или случайности) есть выражение метафизического способа мышления. В диалектическом эволюционном процессе «бетонный мост» Уоддингтона не может быть «жестко детерминированным» мостом. Как следует из рассмотренных нами закономерностей развития, речь может идти лишь о той или иной *степени детерминации,* о соотношении случайных и детерминированных связей, характеризуемых коэффициентом стохастичности G. Например, анализ статистических свойств языка показывает, что параметр G может характеризовать соотношение случайных и детерминированных связей не только между отдельными буквами, но и между словами и фразами, т. е. *на всех структурных уровнях языка.* Точно так же в биологических системах случайные связи существуют одновременно с детерминированными связями не только на низших структурных уровнях («гравий»), но и в тех соединениях, которые образуют следующие структурные уровни и создают нечто единое из «бетонных блоков моста».

Уоддингтон исходит из того, что случайные связи полностью исключают детерминированность. Только при этом условии правомерна упомянутая выше аналогия между случайной изменчивостью организмов и игрой ребенка, набравшего совершенно случайно двадцать строк поэмы Вергилия. Но в том-то и суть, что процесс эволюции вовсе не требует, чтобы случайно возникли сразу целых двадцать вергилиевских строк. Эти двадцать строк уже образованы в течение миллионов лет предшествующей эволюции, а роль мутаций заключается лишь в том, чтобы путем случайного перебора найти не двадцать строк, а всего одну букву, с которой может быть начата следующая строка. Именно такими ничтожно маленькими шагами протекал процесс поиска и закрепления рациональных признаков, причем по мере их накопления образовывалось все большее число детерминированных «правил», которым последующие случайно найденные признаки должны были удовлетворять.

Иной путь тут невозможен: вследствие отсутствия заранее намеченной цели и обобщающих представлений Природа не может использовать для руководства поиском заранее созданный эталон. Отсюда следует, что характерный для осознанного творчества путь от обобщенной цели к частным решениям (дедуктивный метод) для биологической эволюции полностью исключен. Природа не знает заранее, чего она хочет, она просто мутирует «наугад» «букву» за «буквой», а затем испытывает «на прочность» (в смысле приспособляемости и выживаемости) каждый полученный результат. В этом плане поиск в процессе биологической эволюции в большей степени приближается к «слепому»,— однако даже и этот поиск нельзя считать «совершенно слепым». Дело в том, что, созидая новый видовой признак, Природа неукоснительно сохраняет в своей «генетической памяти» весь остальной комплекс признаков, определяющих данный вид.

Для осознанного творчества это условие вовсе не обязательно. В случае необходимости в процессе поэтического творчества можно отказаться от придуманных двадцати строк поэмы и начать поиск «с нуля». Так часто и поступает художник или ученый в тех случаях, когда он чувствует, что предварительно накопленные представления отягощают творческий процесс. **В этом часто заключается още один из важных секретов творческого успеха: суметь вовремя выйти из тупика, наметив обходной путь и новую этапную цель (не теряя при этом цель конечную).** Здесь человеку-творцу обычно приходит на помощь воображение, которым созидающая биологические виды Природа, увы, не наделена. Потому-то Природа и вынуждена созидать длящимися миллиарды лет маленькими шажками, в то время как человек может в один момент озарения «найти точку опоры и перевернуть земной шар».

Подводя итог, можно сказать, что в процессе осознанного творческого поиска «цель-эталон» может возникнуть как «перспективно» (с помощью воображения, интуиции), так и «ретроспективно» (из накопленного опыта и сохраняемых представлений), а в процессе биологической эволюции Природа создает только «ретроспективные» эталоны.

Современные «эвристические» машины действуют скорее по принципу биологической эволюции, нежели по принципам, которым подчиняется творческий процесс у человека. Создавая очередную строку «автопоэмы», машина не формулирует мыслей, заключенных в данной строке. Она «пристраивает» к строке очередное подходящее по грамматическому согласованию и ритмике слово, случайно выбранное из множества других сохраняемых в ее запоминающем устройстве слов. Так возникают удовлетворяющие формальным правилам грамматики и поэтики, но курьезные с семантической точки зрения сочетания слов.

На начальном этапе «электронная эвристика» быда всего лишь очень грубым приближением к процессам эвристического мышления человека. Обусловлено это в первую очередь тем, что электронная машина действовала по рецепту Кэмпбелла, чередуя стадию чисто случайного поиска (определяемую включением генератора случайных данных) со стадией жестко детерминированного отбора. Следующий этап развития «электронной эвристики» включил в себя отказ от локализованного источника стохастичных сигналов и передачу его функций множеству связанных друг с другом логических элементов-ячеек, т. е. переход от «двузначной» логики к «вероятностной», «серой», «нечеткой» логике.

В приведенных сопоставлениях процесса биологической эволюции с процессом творческого мышления человека последний процесс трактовался нами односторонне — в одном лишь интеллектуальном аспекте, без учета его эмоциональных сторон. Между тем хорошо известно, какую большую роль в процессе созидания нового может сыграть, например, творческий экстаз. Подобные эмоциональные факторы пока не поддаются учету. При моделировании эвристического мышления также не удается ввести эти факторы в программы электронных машин. Можно снабдить машину специальными датчиками, заменяющими ей все пять органов чувств человека, но и при этом остается неясным, как закодировать и ввести в программу машины чувства, сопровождающие всякий творческий поиск: разочарование, надежду, досаду, тихую радость или бурный восторг. Этот факт может служить иллюстрацией ограниченности современного кибернетического подхода к анализу сложных взаимосвязанных явлений из области интеллектуальной, эмоциональной и психологической сфер жизни людей. Вместе с тем следует признать, что различные сферы духовной жизни предполагают переработку той или иной информации, и потому к их изучению вполне правомерен единый информационный подход. Отсюда, конечно, не следует, что на данном этапе развития теория информации может предложить пригодный для этих целей в достаточной степени разработанный аппарат.

Мы не касались также вопроса о том, каким образом распределяются функции между сознанием и подсознанием, между корой и подкоркой, полагая, что первая из этих проблем относится к области психологии, а вторая — к области физиологии мозга. С точки зрения теории информации не столь существенно, где «спрятана» детерминированная логика и откуда подается «спонтанный» стохастичный сигнал. На основании имеющихся данных о структуре и механизмах работы мозга можно предполагать, что четкая работа сознания осуществляется за счет детерминированных связей, сосредоточенных преимущественно в коре головного мозга, а спонтанными подсознательными ассоциациями управляют стохастичные сигналы, поступающие в кору из подкорковых областей. Благодаря этим сигналам возникают неожиданные связи между хранящимися в коре функциональными структурами, о которых шла речь раньше.

Можно также предполагать, что открытые физиологией полнообразные процессы возбуждений и торможений определенных областей мозга, связанные с психологическим состоянием индивидуума, в свою очередь, способны стимулировать режим поиска и эвристический процесс путем авторегулирования соотношения случайных и детерминированных связей и приближения его к оптимальному коэффициенту стохастичности G_{opt}.

Анализируя процессы эвристического мышления с точки зрения соотношения формально-логических (алгоритмических) и не подчиняющихся формальной логике (действующих на основе «ослабленных» алгоритмов) процессов, В. В. Бирюков и Е. С. Геллер тоже приходят к выводу о том, что в процессе творческого мышления на каждом структурном уровне (т. е. на каждом уровне обобщений) соблюдается определенное соотношение случайных и детерминированных связей. Данное утверждение, естественно, формулируется в выражениях, несколько отличающихся от принятых в этой книге, но существо самих выводов полностью совпадает с результатами, полученными нами путем анализа соотношений и изменений величин H_o; $H_п$ и G. «В настоящее время,— замечают Б. В. Бирюков и Е. С. Геллер,— становится все более ясным, что интимные процессы мышления человека осуществляются на разных уровнях абстракции, что при этом в работе интеллекта имеет место смена внутренних «языков описания». На каждом уровне абстракции работают как механизмы («алгоритмы») логического преобразования информации, так и схемы эвристического поиска, осуществляется передача информации с одного уровня на другой». Отмеченная передача информации от уровня к уровню есть не что иное, как образование информационных связей между «детерминированными кирпичиками: предыдущего уровня с последующими влияниями вновь образованного высокого уровня на низлежащие уровни т. е. функционирование прямой и обратной связи. При этом случайные связи являются принципиально необходимым дополнением к детерминированным схемам формальной логической дедукции, обладающим «ограниченной «объяснительной силой»... в применении к процессам решения и эвристическим процедурам, присутствующий в мышлении человека».

Естественно, что в силу чрезвычайной сложности процесса эвристического мышления до сих пор этот процесс рассматривался без учета неразрывно с ним связанного развития интеллекта. Во всех рассуждениях о соотношении формально-логических и эвристических операций подразумевался как бы «застывший» на данном структурном уровне интеллект. Между тем в диалектической взаимосвязи явлений два эти процесса не отделимы друг от друга, ибо успешное решение любой конкретной научной проблемы всегда сопровождается развитием бьющегося над решением этой проблемы творческого ума.

Как и всякий эволюционный процесс, развитие творческих способностей осуществляется ступенями, путем перехода с более низкого структурного уровня на более высокую и сложную ступень. Как было показано выше, общей тенденцией развития на каждом структурном уровне является образование детерминированных связей и вытеснение связей случайных, причем предел этой тенденции — полная детерминация. При переходе на новый уровень детерминированные структуры предыдущего уровня становятся своеобразными «кирпичиками» для построения более сложных структур.

Как было показано выше, плодотворность творческого мышления определяется выбором правильных направлений, оптимальным сочетанием в мышлении стохастичности и жесткой детерминации, способностью к переходам с одного понятийного и концептуального уровня на другой, более сложный. (Именно это последнее качество подразумевается в таком, например, выражении, как «способность мыслить другими категориями».) С переходом на новый структурный уровень мышления возобновляется осуществляемый методом «проб и ошибок» процесс упорядочивания («детерминизации») понятий и представлений на новом, более высоком и емком уровне обобщений. На этом уровне процесс продолжается до тех пор, пока интеллект не образует устойчивых связей между вновь освоенными понятиями, превратив их тем самым в «детерминированные кирпичики», из которых можно строить следующую по степени сложности концептуальную систему. Возможность освоения новых, все более высоких структурно-понятийных интеллектуальных уровней мы определяем обычно как «способность к обобщениям», справедливо считая ее одним из главных показателей развития ума. Распространяя эти выводы на «общечеловеческое» мышление, можно обнаружить тенденцию ко все более широкому синтезу, характерную для современного этапа науки. Эта тенденция проявляется в создании таких теоретических направлений, как кибернетика, теория информации, общая теория систем, общая теория консалтинга, социология, бионика и т. п., и может быть определена как познание «вширь». Познание диалектической взаимосвязи и взаимообусловленности явлений на все более высоких структурных уровнях составляет задачу, пожалуй, не менее важную и актуальную, нежели исследование материальных процессов на квантовом уровне, т. е. познание материи «вглубь».

В свете рассмотренных закономерностей развития мысли, приводящих к образованию детерминированных «кирпичиков», создающих условия для перехода на более высокие уровни обобщений, вполне реальной представляется задача разработки специальных психологических тестов, которые позволили бы после соответствующей статистической обработки определить относительное количество детерминированных и нестандартных ответов испытуемых и вычислить для них соответствующие коэффициенты стохастичности G. Усредненный коэффициент стохастичности большого количества заведомо творческих индивидуумов оказался бы близким к оптимальному (G_{opt}). Можно предположить, что величина его будет иметь порядок $G_{\text{opt}} \cong 1/4$, т. е. будет близка к G_{opt} языка. **Мы полагаем, что критерием творческих способностей индивидуума могла бы служить степень приближения присущего ему параметра G к оптимальной величине.**

Следует иметь в виду, что «эвристичность мышления» проявляется не только в сложных творческих актах, порождающих новые шедевры искусства, новые научные теории и т. п. Это важнейшее свойство человеческого сознания обнаруживается и в повседневном общении, и в выборе тех или иных решений, и в конкретных словах и поступках людей. Благодаря этому свойству мышление и поведение человека всегда в той или иной степени непредсказуемо, неожиданно. Отражением «эвристичности мышления» является рассмотренная нами выше стохастичность языка. Так же как и мысль, для выражения которой он предназначен, язык одновременно детерминирован и непредсказуемо спонтанен. Лишив язык детерминации (т. е. выполнив условие $G = \infty$), люди перестали бы понимать друг друга (каждый человек заново изобретал бы новые правила и слова). Если бы язык был абсолютно лишен спонтанности (т. е. при G=0), человеческое общение превратилось бы в обмен общеизвестными истинами и, следовательно, утратило бы всякий смысл.

В спонтанности языка проявляет себя, если можно так выразиться, «повседневная эвристичность мышления» — ведь «маленькие открытия» — это и неожиданно принятое организационное решение, и удачно найденная фраза служебного отчета, и остроумная шутка, и правильно придуманный деловой шаг. Не случайно Луи де Бройль подчеркивал, что французская наука в значительной мере обязана своими успехами тем каламбурам и анекдотам, на которых воспитывался и развивался французский язык. Утверждение это могло бы тоже казаться анекдотичным, если бы оно не заключало в себе глубокий смысл. **<u>Научное открытие всегда «остроумно</u>», так как в его основе лежит неожиданное сопоставление идей или фактов, позволяющее найти в явлениях не обнаруженную до этого времени, а потому и неожиданную для всех связь.**

Из этого, конечно, не следует, что процесс научного или художественного творчества может быть отождествлен с шутливой беседой за общим столом. **Процесс творчества отличается от «бытовой эвристики» прежде всего тем, что требует эрудиции, большой подготовительной работы, заключающейся в освоении и критическом переосмысливании опыта предшественников, сложных методов, понятий и представлений, т. е. требует тех самых детерминированных целенаправленных поисков, отбора и переработки информации,** эвристическую роль которых отверг Д. Кэмпбелл.

Очевидно, что разработка методики оценки «творческого потенциала»— задача отнюдь не простая, поскольку в зависимости от конкретных наклонностей испытуемого необходим тот или иной подход (человек, наделенный творческими способностями к живописи, может быть неспособен к математическому мышлению и т. п.). И все же задачу эту нельзя считать безнадежной, поскольку даже повседневный опыт свидетельствует о том, что всем творческим людям наряду с индивидуальными особенностями присущи и некоторые общие черты. Одной из них является определенная спонтанность их действий, приводящая к тому, что, решая несколько раз одну и ту же задачу, творческий человек (в отличие от человека, склонного к «формализации» своих слов и поступков) будет каждый раз вносить в решение нечто свое. При значительных превышениях G_{opt} эта спонтанность представляет даже определенную опасность.

Другая общая черта индивидуумов, склонных к эвристическому мышлению, заключается в том, что каждый из них априорно допускает возможность ошибочных выводов и потому, получив противоречащие этим выводам сведения, готов искать истоки ошибок и как можно скорее их исправлять. Человека, склонного строить свои концепции по строго формальным правилам, убедить в ошибочности его выводов (даже при очевидной неправомерности исходных посылок) бывает, как правило, много сложней.

Вполне очевидно, что вопрос о *механизме творческого мышления* самым тесным образом связан с *проблемой познания* материального мира, и потому многие выводы, касающиеся соотношения случайных и детерминированных связей в конкретном творческом акте, могут быть распространены и на развитие научного мышления в целом. В этой связи хотелось бы заметить, что именно в интересующем нас аспекте критикуются И. В. Кузнецовым некоторые взгляды Л. Бриллюэна, высказанные последним в его книге «Научная неопределенность и информация». В послесловии к указанной книге, дискутируя с автором, утверждающим, что всякие «измы» лишают свободы творческую мысль, И. В. Кузнецов приходит к тем же выводам о соотношении «свободы» и «ограничений» («внутренней дисциплины») научного мышления, которые выше были сформулированы нами на основе анализа единичного творческого процесса и процесса эволюции языка.

«Можно ли считать подлинно научным, — пишет И. В. Кузнецов,— столь «свободное» мышление, что оно решительно ничем не ограничено, которое, не считаясь ни с какими законами природы, ни с установленными законами логики, с добытыми прежде результатами познания, «творит» все, что ему вздумается? Конечно, нет! **«Свобода мышления» — *действительно научного мышления* — на самом деле предполагает весьма жесткую внутреннюю дисциплину: подчинение определенным законам логики, требованиям доказательности, ясности, последовательности, правилам образования и применения научных понятий и теорий, требованиям соответствия познанным законам природы, соответствия объективной реальности и подчинения строгим критериям истинности.** Конечно, эти требования сами меняются и уточняются в ходе развития познания, некоторые порой даже ломаются совсем. Но «свободная мысль» всегда, даже при наличии в ней таких элементов, выходящих за рамки обычных логических форм, как фантазия, воображение, интуиция, так или иначе согласуется с указанными требованиями и именно поэтому оказывается эффективной. Отрицать это — значит принять всерьез, например, заявление невежды о том, что его «свободное мышление» о переменных математических величинах «ограничено» правилами дифференциального и интегрального исчисления, или сочувствовать утверждениям некоего изобретателя о том, что физика с ее законами сохранения и превращения энергии и вторым началом термодинамики ограничивает «свободу мышления» в создании проектов вечных двигателей первого и второго рода.

Конечно, и математика, и физика, как и все другие науки, благодаря установленным ими законам ограничивают определенным образом «свободу мышления». Но это по существу ограничения от ошибок и путаницы, а не преграды для творчески действующего научного мышления».

Проделанный нами анализ позволяет сделать вывод, что крайняя точка зрения Л. Бриллюэна об отрицательном влиянии всех без исключения «измов» сама тоже является порождением определенного «изма», а именно: метафизического подхода к процессу познания, неизбежно приводящего, как мы уже убедились, к абсолютизации одной из сторон этого диалектически противоречивого процесса (в данном случае — к абсолютизации «научной неопределенности») и к игнорированию другой стороны (определенных «ограничений», обусловленных наличием априорной информации и научных теорий).

Достаточно проследить основные этапы развития взглядов на природу эвристического мышления, чтобы обнаружить удивительное соответствие между господствующей на каждом этапе трактовкой объективного мира и попытками объяснения механизма движения творческой мысли, познающей и отражающей этот мир.

В период расцвета древнегреческой философии, провозгласившей гармонию и единство основных законов объективного мира, возникла гносеологическая теория Платона, утверждавшего, что любые картины внешнего мира являются продуктом «чистого» разума, частными случаями, выводимыми дедуктивным методом из общих идей. В эпоху Возрождения на смену умозрительным построениям схоластов, трактовавшим мир «по Аристотелю», пришла тенденция расчленения сложных явлений и эмпирической проверки любых научных идей. В этот период возникла эвристическая концепция Ф. Бэкона, согласно которой открытия могут быть сделаны не путем дедуктивных логических выводов из созданных разумом общих конструкций, а исключительно путем индуктивного обобщения добытых опытом фактов. Целое тысячелетие на созданную Аристотелем формальную логику смотрели как на главный инструмент доказательства истин. Ф. Бэкон противопоставил дедукции индукцию и сделал попытку превратить индуктивную формальную логику в «логику открытий», позволяющую уверенно отыскивать эти истины. В интерпретации Ф. Бэкона индуктивная логика превратилась в некий универсальный, доступный каждому инструмент для добывания новых истин. Чтобы делать открытия, достаточно лишь, строго следуя установленным логикой правилам, наблюдать и обобщать. Однако индуктивная логика, естественно, не оправдала возложенных на нее Ф. Бэконом надежд. «Расчет на то, что, вооружив ею человечество, удастся добиться невероятного изобилия открытий, оказался тщетным,— замечает М. Г. Ярошевский. — Выяснилось, что индуктивная логика столь же бессильна перед задачей стимуляции и регуляции процессов мышления, как и дедуктивная. Стало очевидно, что правила наблюдений и обобщений, как бы хороши они ни были, так же недостаточны для добывания истины, как и правила построения дедуктивных умозаключений».

Вот почему в противовес детерминистской эвристической концепции Ф. Бэкона К. Бернар высказал иную гипотезу о сущности творческого процесса. К. Бернар утверждал, что в процессе творческих поисков ученый ставит опыты вовсе не для того, чтобы, добыв груду разрозненных фактов, затем извлечь из них методом индукции некий общий закон. **Согласно К. Бернару, опыты нужны для проверки априорно выдвинутых *гипотез.* А что такое гипотеза? Это продукт неожиданных ассоциаций, плод «спонтанного гения», порождение непредсказуемых импровизаций выдающегося ума.**

Нетрудно заметить, что выдвинутая К. Бернаром трактовка эвристики через понятия «непредсказуемости» и «спонтанности» находится в связи с утверждавшимися в это же самое время вероятностными воззрениями на окружающий мир. Но, как это часто бывает в противоречивом процессе развития научной теоретической мысли, отрицание одной крайности породило противоположную крайность: вслед за признанием «спонтанности гения» предпринимались попытки вовсе отвергнуть роль формальной логики в процессе создания и разработки новых научных концепций. В частности, Ф. Шиллер обвинял формальную логику в том, что она произвольно, «соответственно своим предрассудкам» искажает истинные процедуры исследований и научного поиска и потому бессильна регулировать научный прогресс. Параллельно с этим в процессе формирования характерных для этого периода мировоззренческих концепций отказ от жесткого детерминизма Лапласа породил противоположную детерминизму теорию, согласно которой совокупность спонтанных случайных явлений постепенно разрушает присущий миру порядок и в конце концов приведет Вселенную к «тепловой смерти».

Происходящая на наших глазах научно-техническая революция примирила построенные на формальной логике детерминированные программы со спонтанной эвристикой электронных машин. Однако в попытках нового объяснения механизмов творческого мышления на базе эвристического программирования наметился определенный примитивизм. Если в эвристической программе для ЭВМ могут быть четко вычленены во времени и в пространстве жестко детерминированная и стохастичная части, то в сложном потоке человеческой мысли бесполезно искать границы, разделяющие эти диалектически связанные компоненты, как это пытаются делать У. Р. Эшби и Д. Кэмпбелл. В поисках компромисса между стохастичностью и детерминацией мышления У. Р. Эшби и Д. Кэмпбелл остановились на полпути.

Диалектическим решением этой проблемы является признание того факта, что ***неразрывность случайных и детерминированных связей характерна для всех структурных уровней и стадий мышления и для всех этапов развития творческой мысли.***

В свое время А. С. Пушкин очень точно охарактеризовал эту особенность постигающего окружающий мир человеческого сознания: «Не говорите: *иначе нельзя было быть:* коли было бы это правда, то историк был бы астроном, и события жизни человечества были бы предсказаны в календарях, как и затмения солнечные. Но провидение не алгебра. Ум человеческий, по простонародному выражению, не пророк, а угадчик, он видит окружающий общий ход вещей и может выводить из оного глубокие предположения, часто оправданные временем, но невозможно ему предвидеть *случая* — мощного мгновенного орудия провидения». Эти слова А. С. Пушкина свидетельствуют о том, что в решении проблемы соотношения стохастичности и детерминации в окружающем мире интуиция поэта оказалась сильнее, чем детерминистская логика, которой следовал математик Лаплас.

16.7. Причины случайностей

Вопрос о соотношении категорий «причинность» и «случайность» до сих пор остается открытым, так как пути его решения, предложенные различными научными направлениями, нельзя признать удовлетворительными. Так, например, по нашему мнению, не достигают цели попытки введения таких понятий, как «вероятностная (неопределенная) причинность», поскольку между причиной и следствием, по определению, должна предполагаться ***не случайная, а необходимая генетическая связь*** **(причина *порождает* следствие).**

С целью обоснования принципа «вероятностной причинности» Л. Бриллюэн пытается определить границы, разделяющие понятия «причинность» и «детерминизм». Согласно Л. Бриллюэну, *«детерминизм* предполагает «долженствование»: причина должна порождать такое-то и такое-то следствие (и часто добавляется «сразу же»!). *Причинность* принимает утверждение, содержащее «может»: определенная причина может вызвать такие-то и такие-то следствия с некоторыми вероятностями и некоторыми запаздываниями».

Неправомерность предлагаемой Л. Бриллюэном трактовки причинности может быть аргументирована путем следующих логических построений. Допустим, что существует причина C_1, которая, согласно Л. Бриллюэну, способна порождать n различных следствий, имеющих различные вероятности: $p_1, p_2, \ldots, p_k, \ldots, p_n$. Тогда в одном случае по причине C_1 может произойти k-е событие, имеющее вероятность p_k, в другом — любое другое событие из общего числа возможных событий n. Однако при этом возникает естественный вопрос о *причине* того, что в первом случае произошло k -е событие, во втором $(k + 1)$-е событие, в третьем $(k — 1)$-е событие, и т. д. Другими словами, для того, чтобы обосновать принцип причинности с учетом определения, предложенного Л. Бриллюэном, приходится предположить, что между исходной причиной C_1 и n ее следствиями действует случайным образом некая промежуточная, коммутирующая причина C_2.

Разрешением этого противоречия является предположение о том, что C_1 — это не единственная отдельно взятая причина, а целая совокупность (система) различных причин. В таком случае мы приходим к выводу о том, *что* ***между совокупностями множества причин и множества следствий может существовать многозначная вероятностная связь.*** При этом *между каждой отдельно взятой причиной и ее прямым* (и, как справедливо отмечает Л. Бриллюэн, «сразу же порожденным», т. е. не отделенным от исходной причины промежуточными воздействиями) *следствием* необходимо осуществляется ***однозначная, жестко детерминированная связь.***

Отсутствие четкого анализа соотношения объективных вероятностных связей с принципами причинности и детерминизма явилось почвой для попыток опровержения принципа причинности и провозглашения индетерминизма. Примером этого может служить статья М. Бориа «Действительно ли классическая механика детерминистична». Индетерминистским тенденциям противостоит диалектико-материалистическая точка зрения. М. Э. Омельяновский справедливо отмечает: «Квантовая механика признает силовые воздействия на микрообъект и, значит, представляет причинную теорию, так как эти воздействия необходимо порождают соответствующие изменения в дальнейшем ходе движения микрообъектов». Однако для всестороннего обоснования принципа причинности необходимо еще показать, что в этой квантовомеханической форме движения *отсутствуют* такие явления, в которых *нарушается причинно-следственная* связь. А тут вновь возникает нерешенный вопрос об источниках, обусловливающих вероятностный характер движения микрочастиц.

Как отмечает М. Э. Омельяновский, «силовые воздействия представляют причину. Те связи во времени (то есть протекание во времени атомных процессов), которые отражает волновое уравнение, включают в себя также и указанные ***причинные связи***». Это верно, но помимо этих связей, имеющих, несомненно, причинный характер, волновое уравнение включает в себя еще и образующиеся по менее очевидным причинам ***вероятностные связи***. И для того, ***чтобы решать вопрос о соблюдении или нарушении принципа причинности в квантовомеханическом движении, следует анализировать не только силовые воздействия на квантовомеханические формы движения, но и этот выражаемый волновым уравнением Шредингера вероятностный процесс.*** В свете рассмотренных нами энтропийно-информационных соотношений к решению проблемы о соотношении между принципом причинности и случайностью явлений может быть найден новый подход.

До привлечения функции

$$\sum_i p_i \log p_i$$

к анализу энтропийно-информационных отношений понятиями «стохастичность» и «жесткая детерминация» характеризовалось либо наличие, либо полное отсутствие вероятностных внутрисистемных связей, а оценки частичной детерминации (частичной стохастичности) носили чисто качественный характер. Как было показано выше, функция

$$\sum_i p_i \log p_i$$

и вычисленный с ее помощью коэффициент стохастичности G позволяют производить анализ и сравнительную количественную оценку любых переходных стадий детерминации, начиная от полной стохастичности (максимальной энтропии) и вплоть до жесткой детерминации. Поэтому привлечение функции $\sum_i p_i \log p_i$ к анализу вопроса о происхождении и причинах случайных явлений напрашивается само собой.

Нечеткость разграничения объективной и субъективной сущности вероятностных и информационных описаний явлений породила определенную путаницу в терминологии, характеризующей стохастические процессы. Называя один и тот же процесс «хаотическим», «случайным», «стохастическим», «флюктуационным», «спонтанным», «статистическим», «вероятностным», «недетерминированным» «неупорядоченным», «непредсказуемым», «неопределенным», следует иметь в виду, что каждый из перечисленных терминов имеет особый смысловой оттенок. Термины «стохастический», «флюктуационный», «хаотический», «неупорядоченный», «недетерминированный», «спонтанный» характеризуют *объективные свойства процесса*. Термины «вероятностный», «статистический» наряду с объективными свойствами процесса отражают и *методы, с помощью которых исследуется данный процесс.* Термины «неопределенный», «непредсказуемый» выражают исключительно *субъективную оценку процессов или систем* (процесс кем-то непредсказуемый и для кого-то неопределенный).

(С точки зрения семантики термины «детерминированный» и «определенный» равнозначны. Однако в научной терминологии термин «детерминированный» приобрел иной оттенок и применяется для характеристики объективных свойств явлений («детерминированная система», «детерминированный процесс»). Напротив, выражение «определенная система» обычно подразумевает наблюдателя, которым эта система определена. Термин «стохастический» произошел от греческого слова «стохастика», переводимого как «догадка». В современной науке этот термин, так же как и термин «детерминированный», приобрел объективный смысл.

Понятие «детерминация» (кроме его варианта — «жесткая детерминация») обычно так или иначе связывалось с вероятностным описанием, однако до появления теории информации эта связь осознавалась только интуитивно, поскольку не существовало меры, позволяющей производить количественные оценки степени детерминации систем (очевиден лишь тот факт, что жестко детерминированная связь между наступлениями явлений *А* и *В* характеризуется условиями $p_A(B) = 1$ и $p_B(A) = 1$).

Попытаемся уточнить взаимосвязь этих двух понятий, анализируя различные случаи распределения вероятностей в полной группе событий с помощью функции $\sum_i p_i \log p_i$. Напомним,что полной группой в теории вероятностей называют группу событий, удовлетворяющих условию нормировки

$$\sum_i p_i = 1;$$

последнее означает, что в определенные моменты времени, чередуясь в случайном порядке, должны происходить события, составляющие полную группу, кроме них никаких событий происходить не должно. Таким событием является, например, появление в сообщении чередующихся букв из алфавита, имеющего n букв.

Теперь рассмотрим два крайних случая.

Случай 1— жесткая детерминация совокупности n событий, выражающаяся условиями:

$$p_к = 1;$$

$$p_1 = p_2 = \ldots = p_{к-1} = p_{к+1} = \ldots\ p_n = 0;$$

$$H = -\sum_{i=1}^{n} p_i \log p_i = 0.$$

Случай 2 — максимальная стохаотичность совокупности n событий, выражающаяся условиями:

$$p_1 = p_2 = \ldots = p_n = \frac{1}{n};$$

$$H = -\sum_{i=1}^{n} p_i \log p_i = H_{\max}.$$

Простым и наглядным примером максимальной стохастичности совокупности n событий при $n = 2$ является выпадание «орла» и «решетки» при многократном подбрасывании симметричной монеты, для которой $p_1 = p_2 = 1/2$. Противоположным примером жесткой детерминации может служить асимметричный предмет (по типу детской игрушки «Ванька-встанька»), положение которого в силу асимметрии центра тяжести полностью предопределено ($p_1 = 1$, $p_2 = 0$). Во всех промежуточных случаях распределения вероятностей полной группы событий, отличных от рассмотренных случаев 1 и 2 (т. е. когда вероятности $p_1, p_2, \ldots, p_n$ не равны друг другу и ни одна из них не равна 1), совокупность в целом ни жестко детерминирована, ни полностью стохастична (чисто случайна), а соответствуя характеристике ряда ученых, есть одновременно «и той и другой».

Из рассмотрения случая 1 следует вывод о том, что необходимый характер одного из *n* событий полной группы исключает как возможность всех остальных событий, так и стохастичность всей совокупности (это следствие условия $\sum_i p_i = 1$).

Характеризуемая вероятностью p_k степень *случайности* любого (*k*-го) события из группы *n* событий не может быть критерием стохастичности (или, напротив, детерминированности) всей совокупности. Так, например, зная, что наступление события *k* имеет очень малую вероятность p_k нельзя определить с помощью функции $\sum_i p_i \log p_i$ степень детерминированности всей совокупности до тех пор, пока неизвестно распределение вероятностей для всех элементов совокупности $p_1, p_2, \ldots, p_k, \ldots, p_n$. Однако даже по вероятности одного из событий p_k полной группы *n* можно иногда сделать некоторые качественные выводы о *детерминированности* совокупности.

Именно: если $p_k \neq 1/n$, то вероятности группы распределены неравномерно (т. е. не соблюдено условие $p_1 = p_2 = \ldots = p_k = \ldots = p_n$) и, следовательно, совокупность событий обладает какой-то степенью детерминированности. Если p_k велико, можно утверждать априорно, что вся совокупность имеет высокую степень детерминированности, так как в силу условия $\sum_i p_i = 1$ при большом p_k значения

$$p_1, p_2, \ldots, p_{k-1}, p_{k+1}, \ldots, p_n$$

заведомо невелики.

Рассмотренные нами примеры — наглядная иллюстрация известного тезиса, согласно которому ***вероятностные трактовки представляют одну из форм детерминизма.***

Рассмотренные соотношения позволяют сделать еще один важный вывод: они показывают, что само понятие независимости событий в полной группе относительно. В самом деле, независимость событий заключается только в том, что появление события *k* никоим образом не влияет на вероятность событий *k* + 1, *k* + 2 и т. д. (как известно, в теории вероятностей зависимость между событиями *k* и *k* + 1 выражается через условную вероятность $p_k(k + 1)$. Если события *k* и *k* +1 независимы, то $p_k(k + 1) = p(k + 1)$.)

Вместе с тем между всеми событиями полной группы имеется зависимость (связь), выражающаяся условием нормировки

$$\sum_i p_i = 1$$

и проявляющаяся в том, что увеличение (уменьшение) вероятности любого из событий полной группы влечет за собой соответствующее уменьшение (увеличение) суммарной вероятности остальных n — 1 событий.

Другими словами, события полной группы могут быть *независимыми* лишь для случаев *стационарных стохастичных процессов,* т. е. таких процессов, статистические характеристики которых (в частности, вероятности отдельных событий) не зависят от времени. В случае *динамических стохастичных процессов* (а именно такие процессы и интересуют нас в случае исследований эволюционных изменений систем путем оценки изменений количества накапливаемой системой информации с помощью функции

$$\sum_i p_i \log p_i)$$

распределение вероятности p_i изменяется во времени. При изменении распределения вероятностей p_i возникает взаимная зависимость событий, определяемая условием $\sum_i p_i = 1$,и изменяется соотношение степени случайности и детерминированности, определяемое с помощью функции

$$\sum_i p_i \log p_i.$$

Помимо общей взаимосвязи, выраженной условием $\sum_i p_i = 1,$ между событиями, образующими полную группу, могут существовать дополнительные межэлементные связи. Такие связи учитываются с помощью условных вероятностей. Можно отметить четыре характерных случая таких связей.

1) $0 < p_A(B) < 1$, $p_A(B) > p(B)$. Данные условия означают, что между событиями (элементами) *А* и *В* существует в той или иной степени коррелированная стохастичная (случайная) связь ($p_A(B)$ означает вероятность события *В* при условии, что уже совершилось событие *А,* связанное случайным образом (коррелированное) с событием *В.* Степень корреляции зависит от того, какой «силой» обладает знак > в выражении $p_A(B) > p(B)$.)

2) $p_A(B) = 1$. Это условие означает, что между событиями (элементами) *А* и *В* существует *жестко детерминированная связь,*

3) $p_A(B) = p(B)$. Данное условие означает, что связь между *А* и *В* отсутствует *(А* и *В* независимы).

4) $p_A(B) = 0$. При этом условии события (элементы) *А* и *В* несовместимы (появление *А* полностью исключает возможность появления *В).*

На основе рассмотренного соотношения вероятностных и детерминированных связей в полной группе событий можно ввести ряд полезных общих определений.

Мерой случайности (необходимости) единичных событий является их вероятность. Так, например, реализацию события, имеющего вероятность p_1=0,01, можно расценивать как «почти чистую случайность», а реализацию события, вероятность которого составляет p_2 = 0,99, следует считать «почти необходимой».

Мерой стохастичности совокупности *п* событий служит величина

$$H_0 = -\sum_i p_i \log p_i.$$

Мерой детерминированности же совокупности является величина $H_п$, определяемая как

$$H_п = H_{max} — H_o,$$

причем H_{max} и $H_п$ определяются с помощью одной и той же функции $\sum_i p_i \log p_i$, но при разных условиях:

H_{max} — при

$$p_1 = p_2 = \ldots = p_n = \frac{1}{n},$$

H_o — при реальном распределении вероятностей.

Все выводы, относящиеся к совокупности *п* событий, могут быть распространены на системы, обладающие *п* различными признаками (т. е. имеющие «алфавит», состоящий из *п* «букв»).

Критерий детерминированности системы H_o характеризует одновременно и ее упорядоченность, определяемую количеством информации, сохраняемой данной системой, т. е.

$$I_0 = H_п - H_{max} = H_o.$$

Проделанный нами анализ дает основания признать несостоятельность отмеченных Ю. В. Сачковым попыток некоторых авторов соотносить категорию «случайность» только с единичными событиями, а категорию «необходимость» — со множеством событий. ***Понятия «случайность» и «необходимость» — суть общие категории, относящиеся как к единичным явлениям, так и к их***

совокупности. Если вероятность *единичного явления* отлична от единицы, то реализация его будет *случайной.* ***Совокупность взаимосвязанных случайно изменяющихся параметров представляет собой случайный (стохастичный) процесс***. Необходимость единичного события *А* выражается условием *р (А)* = 1. В случайных процессах необходимость проявляется в виде устойчивости усредненных статистических характеристик. Вместе с тем в теории случайных процессов и в теории информации существуют и такие понятия, которые не относятся к единичным событиям, а характеризуют лишь совокупность событий. ***Бессмысленно, например, говорить об энтропии единичного события***. ***Соответственно, только к совокупности (двух и более) событий приложимы такие понятия, как неопределенность, детерминированность, упорядоченность и стохастичность.***

Как справедливо отмечает Н. В. Пилипенко, наука «дает огромный материал для более глубокого понимания с позиций материалистической диалектики вопроса об основаниях, о причинной обусловленности необходимого и случайного».

До сих пор при анализе причин возникновения случайных связей в качестве исходной модели подразумевались некие абстрактные системы или процессы, в которых заведомо преобладают жестко детерминированные связи, нарушение которых происходит в силу каких-то не всегда очевидных причин. Между тем с точки зрения «онтологии» причинно-следственных связей более правомерным оказывается подход к тому же вопросу, так сказать, с другого конца.

В самом деле, ***если нет никаких причин, влияющих на распределение вероятностей тех или иных массовых явлений, вероятности должны быть или равны друг другу, или распределены по Гауссовой кривой.*** Так ведут себя молекулы в замкнутом объеме, если нет нарушающих равновесное состояние газа *причин.* Так чередуются грани подбрасываемой игральной кости, если *нет особых причин* (например, нарушений симметрии), заставляющих, например, грань с пятью очками выпадать чаще грани с одним очком. Так же ведут себя люди, выбирая наугад один из многих предметов (например, лотерейный билет), если *нет причин,* заставляющих отдать предпочтение одному предмету или группе предметов, выделяемых по какому-то признаку из числа остальных.

Приведенные примеры свидетельствуют о том, что при анализе соотношения категорий причинности и случайности с «онтологической» точки зрения более убедительную исходную модель представляет равновероятное распределение («первозданный хаос»), *отклонение* от которого происходит в силу тех или иных объективных *причин.* ***Наиболее общей причиной является воздействие внешней среды,*** приводящее к накоплению информации и к эволюции, выражающейся в дифференцировке значений p_i функции

$$\sum_i p_i \log p_i.$$

В случае полной изоляции от среды мы имеем дело уже не с открытой, а с закрытой системой которая пребывает в состоянии термодинамического равновесия и максимальной энтропии именно в силу *устранения внешних причин.* Такая трактовка природы стохастичных процессов совпадает с точкой зрения П. Лампрехта, называющего ***«негативной случайностью» такую случайность, которая обусловлена не воздействием, а отсутствием тех или иных факторов и причин.***

До сих пор при рассмотрении проблем соотношения категорий случайности и причинности делались попытки найти ***порождающие случайность причины*** и не учитывалась возможность возникновения случайностей, обусловленных ***отсутствием устраняющих их причин***. Этот факт следует расценивать как определенную дань сложившимся в науке традициям, проистекающим из априорного убеждения, что мир жестко детерминирован, а всякое нарушение детерминированных связей возникает только в силу действия случайных причин. На самом же деле вопрос «Почему мир не чисто случаен?» более правомерен, чем неоднократно обсуждавшийся вопрос о том, почему возникает объективная случайность, отвергающая жестко детерминированный лапласовский мир.

Практическое подтверждение этой часто используемой, хотя и не до конца осознанной, методологии — поиски причин несчастных случаев, возникающих на улицах города или в производственных цехах. При анализе такого рода случайностей всегда стараются установить те отсутствуюшие (бездействующие) причины (неполноту существующих инструкций, невыполнение их предписаний и т.п.), которые могли бы предотвратить это событие или, переходя к принятой нами терминологии, более жестко детерминировать процесс, отклонения в котором породили несчастный случай. Очевидно, что при полном отсутствии правил, «детерминирующих» городское движение, на улицах города воцарился бы хаос и каждый перекресток превратился бы в место непрекращающихся катастроф.

При дальнейшем анализе соотношения категорий случайности, необходимости и причинности мы будем исходить из того, что всякое ***массовое явление изначально стохастично***, а *уменьшение* этой *стохаотичности* (детерминация), выражающееся в уменьшении энтропии, происходит в силу тех или иных ***причин***.

Влияние различных причин на стохастичные явления и процессы обнаруживается в изменении их интегральных статистических характеристик: распределения вероятностей, математического ожидания, дисперсии и т. п. Представим себе, что из 1000 бросков монеты «орел» выпал не 500, а 551 раз. Вероятность того, что отклонение это в данной серии опытов произошло случайно, совершенно ничтожна (10^{-3}). Можно уверенно утверждать, что подобное отклонение обусловлено какой-то объективной причиной. Наиболее вероятная причина — асимметрия монеты. Это предположение можно проверить путем тех или иных измерений. Если оно не подтверждается, остается искать причину в различии бросков, например, можно предположить, что бросавший монету путем тщательной тренировки добился того, что «орел» выпадает чаще в тех случаях, когда до броска монета положена вверх «орлом». Как ни сложна подобная тренировка, все же нельзя полностью отрицать возможность такого результата, поскольку при заданных исходной и конечной точках полета выпадание той или иной стороны монеты определяется только **ее исходным положением, местом и силой толчка**. Для стабилизации этих факторов и увеличения вероятности выпадения одной из сторон монеты можно сконструировать специальный механизм. Однако в пределе этот механизм должен быть *абсолютно точным,* так как работа, совершаемая во время переворачивания идеально симметричной монеты, равна нулю, поскольку ее потенциальная энергия в поле земного тяготения одинакова для положения вверх «решеткой» или вверх «орлом». Это значит, что **бесконечно малое увеличение силы броска монеты будет давать ее дополнительный переворот.** Таким образом, мы еще раз приходим к тому же выводу: при отсутствии детерминирующей причины (асимметрии) возникает максимально стохастичный (равновероятный) процесс.

Как было показано выше, детерминированная связь между воздействующей причиной и вероятностными характеристиками явления приводит к изменению величины энтропии, определяемой с помощью функций $\sum_i p_i \log p_i$ или $\int_{-\infty}^{+\infty} W(x) \log W(x)\, dx$. Однако установление той или иной степени детерминации далеко не всегда вскрывает причины, порождающие эту детерминацию. ***Свободное падение тел жестко детерминировано, но в соответствующем уравнении не содержится никаких указаний ни <u>на причину падения</u>***

того или иного тела, ни <u>на причину изменения его скорости</u>. В противовес этому второй закон Ньютона (*F = ma)* говорит о том, что ***сила F является «<u>причиной</u>» движения тела с ускорением а.***

Среди статистических систем и процессов можно встретить такие, когда причина, порождающая отклонения от равномерного или нормального распределения вероятностей, известна априори. Так, например, очевидно, что причина различия вероятностей букв в обычных текстах заключается в грамматических и фонетических правилах языка. Однако сплошь и рядом причина бывает не столь очевидной. Э. Борель приводит следующий пример. Статистикой установлено, что рождаемость мальчиков превышает рождаемость девочек приблизительно на 1%. Этот факт проверен неоднократно, в различное время в различных странах, поэтому есть все основания утверждать, что он причинно обусловлен, но характер этих причин не известен. Можно предполагать и различие условий оплодотворения, и различную выживаемость в эмбриональной стадии для мужского и женского организмов, однако все это — гипотезы, для проверки которых необходимы специальные исследования.

Устойчивое отклонение интегральных статистических характеристик от первоначальных значений всегда происходит в силу тех или иных объективных причин и может служить сигналом к их отысканию. Например, при массовом изготовлении деталей погрешности размеров являются случайными величинами, вероятности которых подчиняются нормальному закону распределения с определенной дисперсией σ. Если на каком-то этапе изготовления и контроля вдруг обнаруживается, что дисперсия изменилась, производство, как правило, приостанавливается с целью выяснения *причины.* Последней может быть ухудшение точности изготовления контролируемых деталей вследствие нарушений регулировки станков, ухудшение точности измерений в результате частичной неисправности измерительного прибора, замена или утомление изготовителя деталей или их контролера и т. п. Множество аналогичных примеров можно встретить в области экономики и социологии, где изменение статистических характеристик, связанных с распределением продукции, финансов, спроса и предложения, с текучестью кадров, динамикой заболеваемости, преступности и т. п., всегда служит сигналом, свидетельствующим о появлении каких-то тормозящих или стимулирующих факторов, т. е. еще не учтенных *причин.*

Заключая сказанное, можно сделать вывод о главной познавательной функции статистических методов: ***они указывают на существование той или иной <u>причины, действующей на процесс</u>*** (хотя в большинстве случаев и не вскрывают ее).

Установление причины или тем более их совокупности, приводящей к тем или иным отклонениям от ***равновероятного или нормального распределений,*** сопряжено, как правило, с большими трудностями. С точки зрения методологии статистических исследований — в тех случаях когда это возможно,— рациональнее ***решать обратную задачу***: ***заведомо определив причину, исследовать методами статистики ее влияния на вероятностный закон***. Если указанным способом единожды установлена связь между причиной и изменением статистических характеристик того или иного процессов, то в случае повторения изменений отыскать причину уже не составит большого труда. Именно такая методология статистических исследований применяется, как правило, в экономической и социологической областях.

Проделанный в этом разделе анализ имел своей целью установить детерминирующие те или иные случайные явления объективные причины. В то же время мы отмечали, что во многих случаях статистические характеристики могут зависеть и от причин субъективных. Так, например, весьма типична ситуация, когда разным событиям и исходам априорно приписываются равные вероятности потому, что объективные статистические характеристики процессов и порождающие их причины не известны. Иногда даже жестко детерминированные связи и строго обусловленные исходы трактуются как случайные только в силу того, что не выявлена объективно существующая связь.

На вопрос «будет ли завтра дождь?» может последовать в разной степени обоснованный ответ. Можно сказать «нет» наугад. Тогда вероятность того, что дождь завтра все-таки будет, составит 0,5. А можно предварительно изучить данные метеослужбы. Тогда ответ «нет» может почти полностью (скажем, на 90%) исключить возможность дождя. Однако и в этом случае нельзя дать полной гарантии, во-первых, потому, что на погоду оказывают влияние объективно стохастические, т. е. однозначно непредсказуемые, факторы, а во-вторых, потому, что даже самая совершенная метеослужба не способна учесть все статистические взаимосвязи огромного комплекса всевозможных причин.

Задача всякого статистического исследования заключается в первую очередь в том, чтобы устранить субъективность априорных оценок и установить вероятностные зависимости, соответствующие объективным связям явлений.

Математическая статистика имеет хорошо разработанный аппарат, позволяющий оценивать достоверность того, что вероятностное описание каких-то конкретных процессов («статистическая гипотеза») соответствует реальному положению дел.

Изначальная стохастичность исследуемых статистической физикой систем обусловлена массовым характером взаимодействий их элементов. Именно этим обстоятельством объясняются рассмотренные А. Я. Хинчиным неудачи многочисленных попыток построения статистической физики на основе анализа жестко детерминированного механического движения молекул с помощью законов Ньютона. Авторы многих работ пришли к выводу о несводимости рассматриваемых статистической физикой массовых молекулярных явлений к законам жесткой детерминации.

Наблюдаемый в опытах дифракции вероятностный характер движения электронов многие ученые тоже склонны были приписывать взаимным влияниям массы этих частиц в электронном луче. Однако эксперименты В. А. Фабриканта, Н. Г. Сушкина, Л. М. Бибермана, проведенные с единичными электронами, пропущенными через кристалл, показали, что даже при устранении взаимодействия электронов характер дифракции не изменяется: множество электронов, пропущенных через кристалл поодиночке, тоже дает на экране кольца, описываемые вероятностной ψ-волной. Ряд физиков данное свойство называют иногда «свободой воли» электронов или других микрочастиц.

Есть мнение, что отклонения эти обусловлены взаимодействием электрона с прибором, с помощью которого экспериментатор пытается зафиксировать импульсы и координаты частиц. Д. И. Блохинцев считает, что круг рассматриваемых взаимодействий надо расширить, включив в него всю «макроскопическую обстановку» (в том числе и упомянутый прибор). Эти поиски причины неопределенности где-то вне самого квантовомеханического движения, по нашему мнению, есть не что иное, как попытка «детерминизировать» микромир, лишить его *изначальной стохастичности,* приписанной ему еще Эпикуром и Лукрецием, стремившимися на этой основе опровергнуть утверждаемый древнегреческими атомистами полный детерминизм и вытекающий из него фатализм. Отрицание детерминизма и фатализма неизбежно приводит к признанию принципиальной неопределенности квантовомеханического движения и к вытекающему из этого выводу о том, что введенное В. Гейзенбергом соотношение неопределенности в принципе отвергает и жесткую детерминированность этого вида движения, и его однозначное описание вследствие невозможности определить одновременно и абсолютно точно и ***координаты и импульсы движущихся частиц.***

Причину «изначальной стохастичности» микромира, по-видимому, надо искать не в «макроскопической обстановке»,— она лежит на более глубоком субквантовом уровне. Не исключено, что на этом уровне действует совокупность каких-то пока еще не исследованных явлений, каждое из которых строго детерминировано, а все вместе они, подобно молекулам максвелловского газа, образуют статистическую систему, имеющую с каждой микрочастицей не детерминированную, а вероятностную связь. Возможно, что именно это взаимодействие с субквантовым уровнем и определяет характер поведения отдельно взятого электрона, проявление его «свободной воли».

Нет сомнения в том, что и макроскопическая обстановка оказывает существенное влияние на статистические закономерности квантовомеханического движения. Но *влиять* — это не то же самое, что *порождать.* Взаимодействие с макроскопическими объектами может влиять только на интегральные статистические характеристики молекулярного или квантовомеханического движения. В квантовой физике влияние макроскопической обстановки выражается в форме граничных условий, накладываемых на волновую функцию, описывающую статистические свойства квантовомеханических систем. Для наглядности можно представить себе, что каждая квантовая частица как бы «зажата» между двумя «слоями».

«Нижний слой» — это субквантовый (микроскопический) уровень. Этот слой является флюктуационно «зыбким» и стохастично «подвижным», чем и обусловлен вероятностный характер движения «лежащих» на этом «слое» частиц. Стохастичность эта не абсолютна: определенная степень детерминированности движения на квантовом уровне выражается в неизменности квантовых чисел, четности, странности, спинов и т. п.

«Верхний слой» — макроскопический уровень. Этот «слой» обладает «жесткостью». Детерминированные связи элементарных частиц с этим «слоем» ограничивают стохастичность их движения: изменения «верхнего слоя» приводят к изменениям интегральных статистических характеристик. Примером такого влияния макроскопического уровня на микропроцессы может служить изменение распределения вероятностей энергий и скоростей молекул газа (или частиц электронного газа) при изменениях параметров среды (давления, температуры). К той же категории взаимодействий относятся и возмущения, вносимые макроскопическими приборами в квантовомеханическое движение микрочастиц.

Что же касается *первоисточника* стохастичного характера движения элементарных частиц, то для его поисков, по-видимому, придется продолжить исследование материи вглубь. Конечно, утверждение о том, что вероятностный характер квантовомеханических явлений зарождается на субквантовом уровне, — это не более чем гипотеза. Однако при всех обстоятельствах останется справедливым следующее утверждение: ***вероятностное поведение микрочастиц обусловлено отсутствием причин, приводящих к жесткой детерминации данного вида движения.*** Открытый на квантовом уровне «изначальный» статистический механизм служит фундаментом, на котором зиждется целая иерархия самых разнообразных стохастичных процессов на более «высоких» уровнях сложности материальных систем.

16.8. Непредсказуемость стохастичых процессов и познаваемость материальных явлений

Признание объективной стохастичности явлений равноценно признанию их принципиальной неопределенности, отсутствия возможности их однозначных предсказаний. В связи с этим некоторые представители различных областей науки высказывают недоумение и тревогу: не означает ли это возрождение агностицизма? Подчас даже высказывается взгляд, что возникающие по неизвестным причинам случайности, которым принадлежит объективиая роль в стохастичных процессах,— это и есть тот установленный самой природой барьер, заглянуть за который науке принципиально нельзя. На самом же деле статистические теории как раз и созданы для того, чтобы преодолевать этот «барьер».

Проводимые методами теории вероятностей и теории информации исследования статистических свойств объективно стохастичных процессов в конечном счете дают возможность осуществлять научные предсказания их хода. Однако предсказания эти, естественно, не однозначны, а даются в вероятностной форме. ***Объективная основа таких предсказаний — отмеченная выше устойчивость усредненных во времени статистических характеристик.*** В этих характеристиках, собственно, и обнаруживается своеобразный «вероятностный детерминизм».

Весьма знаменателен и тот факт, что аппарат теории вероятностей был в основном разработан значительно раньше, чем была осознана объективная природа вероятностных связей. Это — яркая иллюстрация диалектики самого процесса познания.

Вероятностный способ мышления стал развиваться раньше, чем был осознан и признан объективный вероятностный мир.

После признания объективности случайных процессов и рождения кибернетики стало ясно, что без вероятностных связей не могло бы возникнуть осуществляемое по методу проб и ошибок эвристическое мышление, без которого мир невозможно познать. Оказалось, что не признание, а отрицание вероятностных связей порождает агностицизм.

Вероятностное описание обладающих объективной стохастичностью материальных процессов не является приближенным, принятым временно «за неимением лучшего». Оно принято потому, что других способов описания подобных процессов попросту нет. И это вовсе не значит, что эти процессы *непознаваемы или познаваемы только частично.* Напротив, их *вероятностная трактовка полностью адекватна их вероятностной природе,* и не о слабости, а о силе науки свидетельствует тот факт, что для таких сложных явлений она сумела разработать и применить адекватный им вероятностный аппарат.

На современном этапе уже нельзя анализировать гносеологические проблемы не привлекая на помощь идей теории информации и опыт создания обрабатывающих информацию технических средств. Ведь именно эти идеи и средства позволяют науке намного расширить рамки познаваемости материального мира, путем статистической обработки огромного количества данных вскрывать такие закономерности, которые в силу их многоплановости и сложности другими методами познать невозможно.

Как было показано выше, для определения количества информации, сохраняемой тем или иным объектом, необходимо четко разграничивать объективную информацию от информации наблюдений. При этом следует учитывать, что с методологической точки зрения природа системы, регистрирующей результаты наблюдений, не играет принципиальной роли: это может быть и человек, и автомат, и любой созданный природой объект, который в ходе естественной эволюции приобрел способность что-то фиксировать и запоминать. При этом и сам процесс наблюдений может рассматриваться как своего рода эволюционный процесс, в ходе которого, так же как и во всех рассмотренных нами эволюционных процессах, происходит дифференцировка значений p_i функции

$$\sum_i p_i \log p_i.$$

В самом деле, как было показано выше, в начале наблюдений наблюдатель исходит из условия $p_1 = p_2 = \ldots = p_n$. В ходе наблюдений осуществляется «эволюция представлений» наблюдателя об исследуемом объекте, выражающаяся в дифференцировке значений p_i при соответствующем накоплении информации и уменьшении энтропии. В этом аспекте процесс наблюдений можно рассматривать как частный случай взаимодействий, описываемых приведенным выше соотношением:

$$H(X; Y) < H(X) + H(Y),$$

где X — регистрирующая система; Y— объект наблюдений.

Особенность данного случая взаимодействия систем X и Y заключается в том, что в ходе процесса $H(Y)$=const, так как в противовес приведенным ранее утверждениям Л. Бриллюэна энтропия объекта наблюдения $H(Y)$ не зависит от информации, получаемой при наблюдениях. Следовательно, в данном случае:

$$\Delta H(X; Y) = \Delta H = |\Delta I(X)|.$$

Однако известны и такие случаи (например, в квантовой физике, в биологии), когда в процессе наблюдений условия $H(Y)$ = const не выполняются, так как регистрирующее устройство оказывает то или иное воздействие на исследуемый объект.

Введенный и рассмотренный в предыдущих разделах параметр $H_п$, ***характеризующий объективную упорядоченность движения элементов той или иной системы***, в принципе может служить и мерой определенности этой системы для наблюдателя. Однако в большинстве случаев полученная информация об упорядоченности наблюдаемой системы (процесса) не является полной, поэтому определенность системы для наблюдателя ($\Delta I_н$), как правило, меньше, чем ее объективная упорядоченность $H_п$. Как было показано ранее, $\Delta I_н$ становится равным $H_п$ в том случае, когда полностью изучена статистическая упорядоченность системы, т. е. получена максимально возможная информация наблюдений $\Delta I_{н\ max}$, равная информации, объективно сохраняемой системой I_0. Согласно приведенному выражению (15.7) $I_0 = H_п$, следовательно, $\Delta I_{н\ max} = H_п$. При этом количество присущей данной системе *принципиально непредсказуемой информации* определяется как

$$H_o = H_{max} - H_п.$$

Заранее согласимся с тем, что приведенные рассуждения слишком абстрактны, поскольку и без них вполне очевидно, что наблюдатель должен стремиться получить максимальную информацию о структуре исследуемой системы. Вопрос заключается в том, каким образом ее получить. Оказывается, и в этом плане можно сделать некоторые методологические выводы, базирующиеся на опыте разработки и применения предназначенных для обработки информации электронно-вьгчислительных устройств.

С методологической точки зрения весьма важным шагом была разработка так называемых перцептронов — устройств, предназначенных для опознавания образов. Перцептроны являются некоторой приближенной моделью механизмов обработки зрительной информации в анализаторах животных и человека.

В основе процессов опознавания образов с помощью электронных приборов лежит сравнение заимствованных из теории случайных процессов корреляционных функций, получаемых в результате сопоставления развертки изображения опознаваемого предмета с хранимым в памяти его «усредненным образом» (который, в свою очередь, является итогом предварительной корреляционной обриботки признаков множества опознававшихся однородных объектов). В данном случае принципиальное значение ***корреляционного метода заключается в том, что с его помощью удается выделить главные — наиболее характерные, типичные — признаки опознаваемого объекта, исключив признаки второстепенные.***

По-видимому, аналогичный механизм лежит и в основе процессов психического восприятия. Именно благодаря статистической обработке поступающей в мозг информации нам удается установить, что перед нами, скажем, собака, а не теленок, хотя собака эта может быть и овчаркой, и пуделем, и сенбернаром; ***не будь мозг аппаратом определенной статистической обработки данных, мы принимали бы одну и ту же кошку за три разных объекта, в зависимости от того, бежит ли она, сидит или лежит.*** **Не остается сомнения в том, что определение статистических характеристик (распределения вероятностей, корреляции, математического ожидания и т. п.) есть путь выделения наиболее характерных признаков объектов, создания обобщенной картины действительности.** «Отбор существенных переменных не является отказом от точности,— замечает Станислав Лем.—Наоборот, спасая нас от потока несущественной информации, этот отбор позволяет быстрее обнаружить ***целый класс явлений, подобных данному, то есть создать теорию***». Таким образом, ***статистическая обработка множества переменных лежит в самом фундаменте процесса познания, без нее, наверное, немыслимо формирование представлений, понятий, категорий, т. е. всех тех видов обобщений, благодаря которым существует и развивается человеческая мысль.***

Необходимость вероятностных связей в процессе познания заключается также в том, что эти связи позволяют на основании прошлого опыта осуществлять приближенное прогнозирование, создавать «опережающие отображения событий», строить «образы цели», т. е. те научные гипотезы, без которых немыслим познавательный процесс. Именно прогностической силой статистических методов можно объяснить тот факт, что во всех интегративных направлениях современной науки (таких, например, как теория информации, кибернетика, общая теория систем, семиотика и др.) статистические методы исследований играют фундаментальную роль. Да и могло ли быть иначе: ведь взаимодействие случайных и детерминированных связей лежит в самой основе *объективной диалектики* материальных процессов, а одним из важнейших ***инструментов диалектического метода познания этих процессо***в (т. е. диалектики субъективной) служит ***вероятностный аппарат***.

Чтобы представить себе, насколько важна для ***познания материального мира возможность выделения главных признаков на фоне второстепенных с помощью обобщенных статистических характеристик систем, необходимо учесть еще одно важное свойство процессов развития (самоорганизации), которое можно назвать принципом увеличения разнообразия***. Принцип этот можно сформулировать следующим образом.

С переходом на более высокие структурные уровни число «букв», составляющих полный «алфавит» (т. е. число образующих данный структурный уровень элементов — «кирпичиков», имеющих различимые признаки) увеличивается прогрессивно, по закону

$$N_n = N_0^{k^n},$$

где n — порядковый номер рассматриваемого уровня, N_n — число «букв» «алфавита» n-го уровня, k — число элементов («букв»), составляющих одно «слово», N_o — число элементов, составляющих полный «алфавит» уровня, принятого за начало отсчета (n = 0).

Соотношение $N_n = N_0^{k^n}$ отображает принцип увеличения разнообразия признаков, т. е. «прогрессирующий алфавит» систем (следует уточнить, что формула $N_n = N_0^{k^n}$ справедлива при условии, что k не зависит от n, т. е. число элементов «алфавита», составляющих одно полное «слово», на всех уровнях остается неизменным и равным k (каждое «слово» состоит из k букв, каждая «фраза», — из k «слов» и т. д.). Если указанное условие не соблюдается, «прогрессирующий алфавит» выражается несколько иначе:

$$N_1 = N_0^{k_1},\ N_2 = N_1^{k_2} = N_0^{k_1 k_2},\ \ldots,\ N_n = N_0^{\prod_{j=1}^{n} k_j}.$$
)

Наглядным примером действия принципа увеличения разнообразия может служить все существующее и непрерывно пополняемое множество самых разнообразных письменных текстов, которые на нулевом уровне сводятся к алфавиту, насчитывающему около 30 букв (имеются в виду алфавиты различных европейских языков).

«Алфавитом» всей природы служат ограниченный набор элементарных частиц и умещающийся в рамки одной менделеевской таблицы набор атомов химических элементов. На основе этих начальных «алфавитов» строится все присущее материальному миру многообразие веществ и систем. **С этих позиций понятие «алфавит» следует признать универсальным не в меньшей степени, чем понятия «информация», «код», «сигнал» и т. д.**

Очевидно, что увеличение количества различных ***признаков («букв алфавита»)***, происходящее по мере образования более высоких структурных уровней систем, означает одновременно и ***увеличение количества информации***, объективно сохраняемой в структуре этих систем, а следовательно, и увеличение количества информации, получаемой при наблюдении (фиксации) каждого признака из общего числа

$$\sum_{i=1}^{n} N_i.$$

Для оценки порядка величины информации, соответствующей *n*-му структурному уровню, будем считать, что все различимые признаки («буквы алфавита») равновероятны. Как было показано ранее, в этом случае формула Шеннона совпадает с формулой Хартли и количество информации вычисляется как $I = \log N$.

Для *n*-го структурного уровня $I_n = \log N_0^{k^n} =$ $= k^n \log N_0 = k^n I_0$, где $I_0 = \log N_0$ есть среднее количество информации, приходящееся на одну букву (признак) на уровне, принятом за начало отсчета; I_n — среднее количество информации, приходящееся на одну букву (признак) на *n*-м уровне.

Полученное соотношение показывает, что с переходом системы с нулевого на *n*-й структурный уровень количество сохраняемой в ней информации возрастает в k^n раз (нетрудно заметить, что в случае $k \neq \text{const}$ (т. е. когда средние числа «букв» в «слове» на разных уровнях не равны)

$$I_n = \left(\prod_{j=1}^{n} k_j\right) I_0.$$

Данное выражение показывает, что с переходом с нулевого уровня на *n*-й уровень информация вырастает в $\prod_j k_j$ раз, с переходом с уровня $(j — 1)$ на уровень j — в k_j раз).

Именно в этом смысле следует понимать утверждение о том, что ***информация является мерой многоообразия***.

Наименьшим многообразием обладает альтернативный выбор, которому соответствует двоичный код (N_o = 2). На следующем структурном уровне из различных комбинаций k знаков двоичного кода может быть построен алфавит из N_1 букв, причем число $N_1=N^k_o = 2^k$ при больших k может быть сколь угодно велико. Таким образом, именно принципом увеличения многообразия обусловлена неограниченная универсальность двоичного кода и «сводимость» к этому простейшему коду всех существующих в мире алфавитов, систем счисления и других способов кодирования информации.

Пределом специализации ***также является*** ***сведение всякой проблемы, задачи управления, самоорганизации, адаптации к альтернативному выбору (например, нажатию одной кнопки после выбора предварительного решения «нажать — не нажать»).*** Именно поэтому пределом специализации письменного текста оказался двоичный код типа «пауза — А». Простейший двоичный код лежит в основе многих явлений природы (два вида электрических зарядов, два полюса магнитных полей, мембранные эффекты в физиологических процессах, четность и симметрия состояний квантово-механических систем и др.).

Принцип увеличения многообразия диалектически совмещает в себе два противоположных качества материальных процессов: с одной стороны, сводимость самых сложных информационных процессов к простейшим кодам, а с другой — неограниченные возможности усложнения эволюционирующих систем путем перехода на более высокие структурные уровни развития.

Попытаемся оценить некоторые количественные соотношения, характеризующие действие принципа увеличения многообразия, на примере словесного текста. Если принять, что средняя длина слова примерно 6 букв, то количество таких слов, образованных из обычного (точнее, телеграфного) 32-буквенного алфавита, имеет величину порядка $N_1 = 32^6 \cong 10^9$ ($N_o = 32$; n = 1; k = 6). Число «осмысленных» слов составляет около 0,0001 % от общего соличества буквенных комбинаций, т. е. равно приблизителыю $N_1' = 10^5$. Переходя на следующий структурный уровень — на уровень фраз — и принимая по-прежнему k = 6 (т. е. считая, что каждая фраза содержит в среднем 6 слов), получаем $N_2 = 10^{30}$. Если учесть, что количество электронов в составе всего земного шара выражается числом порядка 10^{40}, то становится очевидным, что количество фраз, составленных с помощью 32-буквенного алфавита, практически не имеет границ. Правда, подавляющее большинство этих фраз пришлось бы отбраковать по причине отсутствия смысла или несоблюдения грамматических правил, однако и оставшихся оказывается вполне достаточно для выражения огромного многообразия человеческих мыслей.

Если представить себе, что перед нами не письменный текст, а некая материальная система, имеющая такое же число различимых признаков, как язык (т. е. десятки «букв» на нулевом уровне, 10^9 «слов» на первом уровне и т. д.), не трудно прийти к заключению, что благодаря принципу увеличения многообразия окружающий мир дает нам такое невообразимо громадное многообразие частных признаков, свойств и качеств, что без умения разделять эти признаки на категории, абстрагироваться от несущественных (для данного аспекта) признаков, сопоставлять и обобщать наиболее важные признаки наш разум обречен был бы захлебнуться в этом нескончаемом потоке многообразия.

Принципиальной основой обобщений служит статистический вероятностный анализ, позволяющий выявлять и сравнивать не частные признаки, а усредненные по множеству признаков статистические характеристики. Эту роль вероятностного подхода к явлениям подчеркивал Ф. Энгельс: «До тех пор, пока мы не можем показать, от чего зависит число горошин в стручке, оно остается случайным; а оттого, что нам скажут, что этот факт предусмотрен уже в первоначальном устройстве солнечной системы, мы ни на шаг не подвинемся дальше. Более того: такая наука, которая взялась бы проследить случай с этим отдельным стручком в его каузальном сцеплении со все более отдаленными причинами, была бы уже не наукой, а простой игрой; ибо этот самый стручок имеет еще бесчисленные другие индивидуальные свойства, являющиеся случайными: оттенок цвета, толщину и твердость оболочки, величину горошин, не говоря уже об индивидуальных особенностях, доступных только микроскопу. Таким образом, с одним этим стручком нам пришлось бы проследить уже больше каузальных связей, чем сколько их могли бы изучить все ботаники на свете». Однако даже статистические методы порой оказываются бессильными перед тем нескончаемым многообразием, которое порождает рассмотренный выше принцип. Чтобы убедиться в этом, вернемся еще раз к примеру письменного текста и рассмотрим возможность его анализа с помощью функции $\sum_i p_i \log p_i$. Путем статистической обработки множества текстов определяются реальные величины вероятностей букв ($p_А$, $p_Б$, $p_В$ и т. д.). Поскольку русский алфавит содержит около 30 букв, можно (даже не подвергая тексты статистическому анализу) определить априорно порядок величин вероятностей: примерно 1/30, т. е. около 0,03. В действительности вероятности эти колеблются от максимального значения $p_{«с»}$=0,11 до минимального значения $p_{«ф»}$= 0,002. Подставив реальные значения p_i в формулу $\sum_i p_i \log p_i$ (где $i = А, Б, \ldots, Я$), получим реальную энтропию словесного текста.

Предположим теперь, что мы хотим повторить этот анализ на следующем структурном уровне — на уровне слов. Как было отмечено выше, число одних лишь осмысленных слов составляет величину порядка $N_1' = 10^5$. Значит, вероятность появления одного какого-то слова (согласно нашей терминологии слово теперь становится «буквой» присущего данному уровню «алфавита») выражается величиной порядка $p_i \simeq 10^{-5}$. Чтобы установить реальную вероятность какого-то слова, пришлось бы статистически обрабатывать тексты, состоящие из миллионов слов. Начиная с уровня n = 1 порядок величин вероятностей, определяемый как $p_i \cong 1/N_0^{k^n}$ (или, в более общем случае, как $p_i \cong \frac{1}{N_0^{\prod_{j=1}^{n} k_j}}$), с переходом на каждый следующий уровень уменьшается столь стремительно, что операции с этими величинами теряют какой бы то ни было практический смысл. А раз неизвестны p_i, значит, нельзя привлечь для количественного анализа функцию $\sum_i p_i \log p_i$. Этим в первую очередь определяется трудность подсчета количества информации, сохраняемой в структуре сложных многофункциональных систем.

Один из возможных путей преодоления этой трудности заключается в следующем. Влияние связей на высоком структурном уровне (например, связей между словами) учитывается как взаимная корреляция элементов более низкого уровня (в данном примере — как взаимная корреляция букв). Затем на нижнем уровне (на уровне букв) используется функция $\sum_i p_i \log p_i$, причем количество информации определяется с учетом обусловленной корреляцией взаимной зависимости условных вероятностей p_i. К сожалению, далеко не для всякой реальной системы удается учесть все корреляции, а затем перевести их в биты. Несомненно, что отмеченные методологические трудности надо расценивать лишь как болезни роста, которые наука сумеет со временем так или иначе преодолеть. Что касается перспектив развития статистических и информационных методов, то здесь можно с полной уверенностью утверждать, что в процессе познания материального мира их роль будет и впредь возрастать неизменно по мере того, как научное и общечеловеческое мышление будет подниматься на все более высокие уровни обобщений. Этот вывод является одним из следствий принципа увеличения многообразия, согласно которому многообразие признаков, а следовательно, и потребности привлечения вероятностных корреляционных методов, позволяющих выделить главные признаки на фоне второстепенных,— непрерывно растут в соответствии с формулой $N_n = N_0^{kn}$.

ЗАКЛЮЧЕНИЕ

Главный вывод, который может быть сделан на основе многих обсуждавшихся в этой работе положений и фактов, можно кратко сформулировать следующим образом:

в решении извечного спора о соотношении случайных и жестко детерминированных связей в природе <u>момент привлечения теории информации</u> в известном смысле является <u>ключевым</u>.

Для решения спора недостаточно было не только понятий о случайных и необходимых событиях, но и их количественных оценок с помощью «обычного» вероятностного аппарата. Для этого требуется более ***сложная интегральная количественная мера***.

До рождения теории информации статистическая формула энтропии использовалась очень односторонне, в связи со Вторым началом термодинамики. Было показано, что состояние термодинамического равновесия, к которому стремятся все предоставленные самим себе («закрытые») физические системы, есть не что иное, как равенство вероятностей всех событий *(p_i)*, которому соответствует максимальная энтропия. Заслуга создателей теории информации, и в первую очередь К. Шеннона, заключается в том, что функция $\sum_i p_i \log p_i$ начала применяться для оценки систем, вовсе не стремящихся к термодинамическому равновесию, а, напротив, благополучно пребывающих в состоянии не с равными, а с дифференцированными значениями p_t. Именно степень отличия от состояния равномерного распределения вероятностей и оценивается количеством информации, сохраняемой (а следовательно, существующей объективно) в структуре сложно организованных систем.

В работе было показано, что с помощью количественной меры информации можно производить оценки структурных характеристик в состоянии статистического равновесия (гомеостазиса) заранее сформировавшихся сложных систем. Более важно, однако что функцию $\sum_i p_i \log p_i$ можно использовать для анализа эволюционной динамики развивающихся систем. Выводы, полученные в результате такого подхода, оказываются общими в той же мере, в какой понятие информации применимо ко всем видам эволюционирующих систем. А замечательная функция $\sum_i p_i \log p_i$ становится именно той формулой, с помощью которой «демон Лапласа» может объять прошедшие, настоящие и будущие состояния мира, с той лишь существенной оговоркой, что перед умственным взором предстает не запрограммированный метафизический, а непредсказуемо изменяющийся диалектический мир.

Вывод о том, что функция $\sum_i p_i \log p_i$ может служить обобщенной математической моделью всех протекающих в мире эволюционных процессов, является результатом большого пути развития теоретико-информационных, кибернетических и теоретико-познавательных идей. С помощью этой функции Л. Больцман и К. Шеннон решали, казалось бы, совершенно различные по своей сути задачи. Но по мере осознания связи информации и энтропии становилось все более очевидным, что на самом деле и Л. Больцман и К. Шеннон исследовали единые закономерности энтропийно-информационных процессов, хотя и подходили к ним с противоположных сторон. Оказалось, что трансформация функции $\sum_i p_i \log p_i$ при переходе от рассматривавшегося Л. Больцманом состояния термодинамического равновесия к изучавшимся К. Шенноном и его последователями информационным системам — это и есть описанный лаконичным языком математики обобщенный эволюционный процесс. Осознание этого факта позволяет использовать функцию $\sum_i p_i \log p_i$ подобно любой другой математической модели: в чисто математическом аспекте исследовать ее абстрактные свойства, а полученные таким образом выводы интерпретировать применительно к эволюционным процессам различных конкретных систем.

Подчеркнем еще раз, что анализ эволюционных процессов с помощью функции $\sum_i p_i \log p_i$ позволил обнаружить тенденцию перераспределениявероятностей p_i, приводящего к постепенному переходу от предельно стохастического состояния системы (при котором все p_i равны друг другу) к детерминированному состоянию с максимально дифференцированными значениями p_i. При таком рассмотрении соотношения случайных и детерминированных связей выявляется диалектическое единство противоположностей этих связей, которые становится возможным рассматривать не как мертвые и застывшие, а как «живые, условные, подвижные, превращающиеся одна в другую». Такой подход может быть с полным правом противопоставлен метафизическому взгляду с присущим ему стремлением противопоставления необходимых и случайных явлений.

Как было показано, метафизический метод тоже эволюционировал. Если раньше противопоставление детерминации и случайности имело характер категоричный, то современные «метафизики» допускают тут некоторый компромисс. Категоричность прежних метафизических воззрений выражалась в стремлении доказать, что случайные и необходимые связи не имеют права сосуществовать. Иную метафизическую концепцию обнаруживаем мы в трудах некоторых современных ученых. Казалось бы, отдана дань и роли связей случайных, и роли связей детерминированных, но... воспитанная веками метафизическая традиция все-таки не позволяет допустить совмещение в одном процессе двух противоположных начал. Отсюда и вытекают некоторые неубедительные выводы, анализ которых дан в настоящей работе. Можно надеяться, что проведенный в этой книге анализ общих свойств эволюционных процессов, базирующийся на идеях, методах и результатах теории информации, послужит достаточно веским опровержением высказанной Уоддингтоном пессимистической точки зрения на эффективность информационного подхода к решению биологических проблем. Одна из основных задач книги как раз и заключалась в обосновании мысли о том, что внедрение методов теории информации в биологию — это не простая перефразировка старых идей, а новый образ мышления (его можно назвать «информационным мышлением»), который, несмотря на ряд методологических трудностей, сулит принести немало плодов.

Есть основания утверждать, что с рождением теории информации завершился очередной виток спирали в развитии вероятностного подхода к миру, начало которому положили в свое время Людвиг Больцман и Джемс Максвелл. На предыдущем витке аналогичную роль сыграли Гераклит, Эмпедокл, Аристотель и Эпикур. Трансформация научного мировоззрения от детерминистического фатализма Демокрита до диалектической трактовки взаимосвязи случайного и необходимого у Эпикура аналогична переходу от лапласовского детерминизма к современной вероятностной трактовке объективных явлений.Установление единой меры количества информации, пригодной для оценки упорядоченности движения, существующего в материальных системах всех степеней сложности, служит еще одним подтверждением позиции материалистического монизма. Единый информационный подход к анализу процессов диалектического развития показывает, что способность к самодвижению (или — в кибернетической терминологии — самоорганизации) в той или иной степени присуща всем материальным системам начиная с объектов неорганической природы и кончая самыми сложными процессами, связанными с жизнью людей.

Интерпретация информации как единой статистической меры упорядоченности движения оказывается настолько общей, что включает в себя как частный случай и понимание информации, заимствованное из обихода. Именно такой вид информации определяет упорядоченность движений, происходящих в социальной, экономической и научной областях. Проведенный анализ взаимодействия случайных и детерминированных связей в эволюционных процессах позволяет конкретизировать некоторые общие положения диалектики. Представим это в виде таблицы:

Принципы диалектики	Их конкретизация в терминах теории информации и теории случайных процессов
Развитие как движение простого к сложному по восходящей линии	Развитие как процесс прогрессивного накопления информации, определяемой условиями: $\frac{\partial^2 I}{\partial N^2} > 0; \quad \frac{\partial^2 I}{\partial M^2} > 0$
Развитие как результат борьбы противоположностей	Развитие как результат «борьбы» между стремлением системы к эволюционной гибкости и к сохранению целостности, выраженной в форме противоречия между увеличением детерминированных связей л сохранением связей случайных
Единство противоположностей в процессах развития	Стремление к оптимальному соотношению случайных и детерминированных связей
Переходы количества в качество (качественные скачки)	Накопление определенного количества информации и последующие переходы на более высокие структурные уровни развития.
Развитие как движение по спирали с повторением на каждом более высоком витке основных этапов предыдущего витка	Продолжение развития системы на более высоком, структурном уровне с сохранением тенденции накопления информации, детерминации и уменьшения $G=H_O/H_H$

Рождение теории информации и системно-информационного подхода к изучению сложных материальных объектов — это важный шаг в методологии научного исследования. Но не следует преувеличивать достигнутые в этом плане успехи. Дело в том, что сам структурно-информационный подход к исследованию сложных объектов приносит не только новые знания, он раскрывает также и то, насколько поддающиеся математической формализации упрощенные модели еще далеки от реальной сложности порождаемых природой систем. Тем более возрастает роль методологических обобщений. Опыт одного из таких обобщений и является темой данной книги. На его основе делается попытка анализа общих закономерностей процессов эволюции, адаптации и самоорганизации развивающихся систем.

Приложение

Альтернативный подход к теории информации

(Ниже приводится альтернативный подход к теории информации, который разделяет рад ученых в области информациологии, в частности, академик И.И.Юзвишин, его последователи и ученики).

На современном этапе информация становится главным ресурсом научно-технического и социально экономического развития мирового сообщества. Информация - это всеобщий, бесконечный, единый законопроцесс фундаментальных отношений, связей взаимодействий энергии (Э), движения (Д), массы (М) и антимассы (А) микро- и макроструктур Вселенной.

Вселенная – информационна. Мир - информационен. Первичное - информация, вторичное - материя. Информация - это поистине неисчерпаемый ресурс мирового сообщества. Информация не только существенно влияет на ускоренное развитие науки, техники и различных отраслей народного хозяйства, но играет определяющую роль в процессе воспитания, образования экономических и хозяйственных отношений, культурном общении и связи, как отдельных людей, так и коллективов, городов, стран.

В настоящее время в науке наблюдается определенная пауза, спад открытий. Этот период непременно вызовет этап пересмотра, анализа и критики существующего научного наследия. Наступает период пересмотра основ научных воззрений. Такой работой является книга академика И.И.Юзвишина «Информациология». Особенностью монографии является принципиальная новизна в подходе к рассмотрению и обоснованию научных основ мироздания.

Рассматривая фундаментальные законы Вселенной, макро- и микромир, процессы материализации и дематериализации, структуризации и деструктуризации с точки зрения информационного подхода, автор монографии приходит к ряду принципиально новых положений и выводов.

С учетом изложенного представляется актуальным анализ и исследование научных основ информациологии.

Цель альтернативного подхода к теории информации состоит в том, чтобы показать научную обоснованность, истинность и корректность основных положений, методов и форм информациологии.

Информациология является фундаментальной наукой, обьединяющей все науки. Информация являеется основной и безальтернативной формой проявления законов и явлений мира, подчинена собственным наиболее общим законам. Опираясь на изложенные основные положения и методы можно расширить поле категорий, определений, законов информации, определить краевые условия информационных областей, критерии метерализации и дематериализации, показать действие организующих и формирующих факторов энергетико-полевых и информационных условий. На примере единиц измерения основных физических величин показать, что все единицы измерения всех наук могут быть приведены к единой унифицированной единице измерения носящей в себе информационное начало.

В рассматриваемом подходе применены методы теории множеств, метрических пространств, методы оптимизации, статистически-стохастические методы и методы определения краевых условий физических областей.

В излагаемом приложении проведен анализ и исследование научных основ информациологии. Результаты могут быть использованны, при изучении информациологии, ее основополагающих категорий, методов и принципов, в определении общих фундаментальных основ всех наук. Полученные результаты и выводы могут иметь определяющее гносеологическое значение в философском осмыслении окружающего мира. Ряд законов могут быть использованны в генной инженерии, теории множеств и метрических пространств.

Суть всего сущего есть информация и есть Бог - как духовное начало. В Евангелии от Иоанна в первых строках сказано: "В начале было Слово /Логос/, и слово было в Бога. Оно было в Бога из поконвеков. Всё через него возникло, и ничто, что возникло, не возникло без Него". Таким образом, изначально определяется Логос как суть сущего, как информация. Бог как Дух, творящая и созидающая основа проявления информации. Информация - это фундаментальный субстрат мироздания. Информация первична, материя вторична. Платон определял основой всего сущего идею, т.е. информацию. Гегель в работе "Наука Логики" определял основой всего сущего абсолютную идею, которая отчуждает себя в различных формах - как материальный процесс. Информация - это генерализационно-единая фундаментальная основа всех процессов и явлений, происходящих в микро- и макродинамических структурах и представляется как распределенно-локальное информационно-сотовое (материализованное и дематериализованное) самоуправляемое поле Вселенной, первоосновой которой является фундаментальный микро- и макромерный автоинформгенезис, постоянно обеспечивающий непрерывные процессы кодирования, декодирования, космической автогенерации и электромагнитно-резонансного равновесия всех информационно-кодовых структур единого распределенного информационно-сотового пространства. Информация - это генерализационный безначально-бесконечный единый законопроцесс микро- и макромерных отношений, взаимосвязей и взаимосохранения энергии движения и массы, на основе автоосцилляционной. резонансно-сотовой, частотно-квантовом и волновой природы света, тепла, звука и других свойств и форм в микро- и макроструктурах Вселенной. Информациология - это генерализационная наука о всех информационных явлениях, микро- и макродинамических процессах беспредельной Вселенной. Объектом информациологии является объективная реальность безначально-бесконечной информационной Вселенной, существующая независимо от нашего сознания внутри нас, вне нас, между нами и вокруг нас - везде и всюду выступающая как объект нашего существования и познания. Предметом информациологии являются исследования информационных микро- и макродинамических процессов, происходящих во Вселенной во взаимоотношениях, во взаимосвязи и во взаимодействии с овеществленными и неовеществленными атрибутами материализации и дематериализации, источниками аннигиляции и автогенерации, а также процессы рецепции, передачи, хранения, обработки, визуализации и познания информации. Методом познания и отражения информации, как предмета исследования является практика — критерий истины, теоретические исследования и расчеты.

Информационные поля имеют многоуровневую структуру. Самым глубоким уровнем является тот уровень информации, который имеет вечную реальность и через который выражаются все уровни. Этот уровень в полевом определении выступает как информационный континуум - безначально-бесконечное беспредельное пространство информации. Единица идеального отношения состояния определенности называется идеальным ***субинформационом***.

По классификации информациологии И.И.Юзвишина размер идеального субинформациона 10^{-100}. То что на физическом уровне может быть воспринято как протопричина, протовещество неуничтожимое даже в процессе Большого Взрыва - это информационное поле идеальных, гипотетических субинформационов. С точки зрения информации это простейшие составные элементы информационного мира. Скорость распространения субинформационов бесконечна.

Единицу элементарного отношения обладающую энергетическим квантом можно представить как ***информацион***. Определение его такое:

квант поликорреляций в материализованном и дематериализованном информационном поле пространства Вселенной, обеспечивающей квант энергодействия $i=h\nu$. где i — фундаментальная информационная коварианта $10^{-\infty} < i < 10^{\infty}$, $i<10^{-34}$ Дж.с. Информационы пронизывают как нейтрино все материализованные и дематериализованные участки пространства Вселенной своим несчетным количеством. В кубическом сантиметре пространства

$$\forall \sum_{i=\infty} i_i \exists Э = \int_V \rho d v c^2 = 10^{\infty} \text{ Дж} \times \text{см}^{-3}$$

$$\text{Inf } i << 6{,}62 \times 10^{-34} \text{ Дж.с.}$$

$$6{,}62 \times 10^{-34} << \text{Sup} i \leq 10^{\infty} \text{ Дж.с.}$$

Элементарный информацион — это квант симметричных отношений информационного поля, обладающий нулевой массой и импульсом симметричного самоотношения $m_0 \frac{\partial i}{\partial t}$, собственной степенью свободы и скоростью превышающей скорость света. ***Информациор*** это спинорно-сферическое поликорреляционное отображение глубинно-тончайших информационных процессов вакуумных, материально-полевых отношений. ***Информациоры — это множества дипольных полей***. Они лежат в основе корреляционных взаимодействий распространяемых в материализованном мире практически с очень огромной скоростью. Информациорные преобразования позволяют определить время отклика информационно-сотовой сети при переходе от одной ее структуры к другой. ***Информациал*** — это структурное образование, состоящее из сочетания иформационов и информациоров, обладающее энергетико-полевыми свойствами и однозначно определенным отношением к окружающему информационному полю. ***Простой информациал*** — это эргетико-полевой информациал, имеющий следующие виды:

$$E=kT;\ E=mc^2;\ E=h\nu;\ 4\pi e;\ 4\pi GM.$$

Поскольку поле протоиинформации или субинформации подчиняются закону всеобщей глубинной структуризации, кодированию и сотовости, то образуются различные отношения взаимообусловленности и опосредствования и данная область информации может выступать в виде информационной структуры

Вне пределов физических полей, когда калибровочная размерность объектов протоиинформации достигает значения, находящегося за пределами кванта энергодействия *h* имеет место информационно-полевой информациал $10^{-100} < iH < 10^{-34}$.

Так, как данный информациал определен только в пределах информационных полей, то его можно назвать субинформациалом.

Он является движущим фактором и организации дематериализованных информационных структур. Составной информициал — это сложное структурное образование, имеющие общее структурное выражение вида $I = \bigcup_{i=1}^{\infty} h_i v_i \sum_{n=1}^{\infty} I_n$.

Составной информациал может выражать как корреляционные отношения, так и некорреляционные. Примером информациалов выступают все элементы таблицы Менделеева и все их сочетания, которые являют собой многообразие неорганических и органических соединений. Корреляционные отношения проявляются например у молекул между атомами, некорреляционные между отдельным атомом и, например, Луной. В отношениях структурности нет замкнутых циклов информационных уровней. Так, строение атома и Солнечной системы имеет некоторое внешнее подобие, но это не тождественность. Существует процесс становления и созидания информационных структур и их разрушения - процесс материализации и дематериализации или процесс структуризации и деструктуризации. ***Процесс структуризации*** — это причинно-следственный процесс, в ходе которого происходит реализация более сложных информационных отношений. Информация упаковывается, укомплектовывается в глубину и вширь. ***Процесс материализации*** — это процесс создания материальных структур. ***Процесс дематериализации*** — это процесс развала материальных структур. Эти два процесса в общем случае только для материализованной информации совпадают с процессами структуризации и деструктуризации информации, поскольку процесс структуризации охватывает еще бесконечно-безначальную область идеальной информации, которая непредставима в материализованном мире, но составляет его глубинную сущность, как информационнная основа. Для дематериализованных информационных полей вводится понятие составного субинформациала, который имеет вид: $I_c = \sum_{i=1}^{\infty} H_i v_i \bigcup_{n=1}^{\infty} I_n$, где H_i, v_i - организующий фактор информационных структур, $\bigcup_{n=1}^{\infty} I_n$ — различные комбинации и сочетания информационных структур, v_i — информационно связная частота.

<u>Плотность информации в данной локальной области - это количество информационов в единице объема</u>

$$\rho_a = \frac{\sum_{j=1}^{\infty} i_j}{V_a}$$

Плотность информаиила - это количество информаиионов в единице объёма данной структуры

$$\rho_{\alpha_i} = \frac{\sum_{i=1}^{\infty} h_i v_i \bigcup_{i=1}^{\infty} I_i}{V_\alpha}$$

Плотность идеальных информаиионов - это количество идеальных информаиионов в единице объёма. Плотность идеальных информационов максимальна в момент нескольких мгновений после Большого Взрыва.

1. Закон сохранения информации.

Суммарное количество информации I и энтропии S_i-го состояния пространства его локальной области при любых информационных процессах остается всегда постоянной.

$$\sum I_i+\sum (S_i)=\text{const}.$$

2. Генерализационный закон сохранения информации.

$$\Delta I = \sum_{i=1}^{\infty} I_i = \text{const}$$

3. Закон информационного равновесия Вселенной.

Все тела и процессы Вселенной находятся в информационном равновесии в силу постоянных безначально-бесконечных локализованных и делокализованных отношений и взаимосвязей.

4. Закон постоянного изменения информации.

Все тела, молекулы, атомы, информационы полей и пространства находятся в постоянном взаимоотношении, взаимосвязи, взаимозависимоепш, взаимодействии и взаимопревращении, обеспечивая тем самым постоянное движение и изменение Вселенной.

5. Закон структурированности

Информационное поле Вселенной структурировано на всю глубину и во всех пределах.

6. Закон формирования структур.

При понижении уровня энергетических информациалов степень структурированности информации в данной локальной области повышается, при повышении - падает.

7. Генерализационный закон информации.

Информация — это генерализационно-фундаментальная субстанция единого кодово-сотового пространства Вселенной, включающего воздух, воду, землю, солнечные и другие светоносные лучи, поля, их следы и весь спектр космических излучений материапизованных и дематериализованных сред и выражается через массу, скорость, энергию и другие формы проявляющиеся в процессе материализации и дематериализации.

Проведем исследование генерализационного закона информации. Процесс структуризации и деструктуризации, материализации и дематериализации всегда происходит с какой-то определенной скоростью, которая явным или неявным образом зависит от температуры, а значит от энергии. Информациология, таким образом, устанавливает на информационной основе единство между: информацией, энергией, движением, массой, пространством и временем вводя логико-символьные записи структурных образований или составных информационалов. Так, кванторное выражение

$$\forall \bigcup_{i,j}^{\infty} I_i^j \exists \bigcup_{i,j}^{\infty} Э_i^j \bigcup_{i,j}^{\infty} Д_i^j \bigcup_{i,j}^{\infty} М_i^j \bigcup_{i,j}^{\infty} П_i^j \bigcup_{i,j}^{\infty} t_i^j$$

где I - универсальное множество (информация), J-ый процесс i-го информационного поля, а также известного соотношения $Э=mc^2$. Поскольку в неявном виде $I \sim F(Э, Д, М, П, t)$

$$Э = mc^2 \sim 1$$

Учитывая безначально-бесконечные процессы материализации и дематериализации, то можно представить формулу информации в следующем виде:

$$I = \xi \dot{t} x_i \chi_i \dot{\beta},$$

где ξ- коэффициент пропорциональности, i - информационная ковариантa, x_i - i-ый объект, χ_i - скорость объекта, $\dot{\beta}$ - скорость процесса материализации и дематериализации.

<u>Исследование:</u> Покажем, что кванторное выражение $\forall \bigcup_{i,j}^{\infty} I_i^j$ - существует группа кванторных объединений $\bigcup_{i,j}^{\infty} Э_i^j \bigcup_{i,j}^{\infty} Д_i^j \bigcup_{i,j}^{\infty} М_i^j \bigcup_{i,j}^{\infty} П_i^j \bigcup_{i,j}^{\infty} t_i^j$ которую можно представить в виде

$$I = \xi \dot{t} x_i \chi_i \dot{\beta}.$$

Будем исходить из выражения $I = \xi \dot{t} x_i \chi_i \dot{\beta}$. Поскольку χ_i -скорость объекта, то имеет место представление $\chi_i \sim \frac{\partial x_i}{\partial t}$, $\dot{\beta}$ - скорость процесса материализации, то она всегда явным или неявным образом зависит от количества движения $\dot{\beta} \sim m\vartheta \sim m\frac{\partial x_i}{\partial t}$.

Выражение $\xi \dot{t} x_i$ - есть по сути дела координатно-информационным переводным маркером или обобщенным коэффициентом перевода α_i, который принимает участие в отношениях равенства, но отсутствует в выражениях пропорциональности. Тогда будем иметь:

$$I = \alpha_i m \frac{\partial x_i}{\partial t} \cdot \frac{\partial x_i}{\partial t} = \alpha_i m\left(\frac{\partial x_i}{\partial t}\right) \sim m\vartheta^2 \sim mc^2$$

С другой стороны из выражения

$$\bigcup_{i,j}^{\infty} Э_i^j; \bigcup_{i,j}^{\infty} Д_i^j; \bigcup_{i,j}^{\infty} М_i^j; \bigcup_{i,j}^{\infty} П_i^j; \bigcup_{i,j}^{\infty} t_i^j,$$

группа $\bigcup_{i,j}^{\infty} П_i^j; \bigcup_{i,j}^{\infty} t_i^j$, наряду с конструкциями $\partial x_i \bullet \partial t$ содержит $\frac{\partial x_i}{\partial t}$.

Также $\bigcup_{i,j}^{\infty} М_i^j$ -содержит m_i. $\bigcup_{i,j}^{\infty} Д_i^j$ - содержит $\frac{\partial x_i}{\partial t}$. $\bigcup_{i,j}^{\infty} Э_i^j$ - содержит $m\left(\frac{\partial x_i}{\partial t}\right)^2$.

Тогда в более явном виде можно записать следующие отношения:

$$I \sim \bigcup_{i,j}^{\infty} m_i \left(\frac{\partial x_i}{\partial t}\right)^2 . \bigcup_{i,j}^{\infty} m_i \frac{\partial x_i}{\partial t} \cdot \frac{\partial x_i}{\partial t} = \bigcup_{i,j}^{\infty} m_i \left(\frac{\partial x_i}{\partial t}\right)^2 \sim mc^2 \sim m\vartheta^2$$

Что доказывает эквивалентность данных форм представления.

8. Закон информационного единства.

В любой локализованной области α *пространства Вселенной всегда имеется микро- или макрообъект, материализованное или дематериализованное поле, его след* X_i *для которого существуют объекты* Y_j*, имеющие в этой области информационные отношения* $R_k X_i$.

$$\forall_{i=1}^{\infty} X_i \exists_{j=1}^{\infty} Y_j R X_i$$

По сути дела, это есть закон единого информационного детерминизма для любого информационного объекта. Это значит, что в информационном поле не существует замкнутых и неопределенных областей. Именно поэтому можно говорить, что информационное пространство познаваемо во всем объеме и на всю глубину. Оно открыто относительно процесса познания. Именно сейчас можно ответить на вопрос Эйнштейна — почему мир понятен? — Мир понятен потому, что информационное поле не замкнуто. Анализ Всемирного закона тяготения на основе принципов информационного подхода позволяет рассмотреть этот закон и ответить не только на вопрос как это происходит, но и дать ответ на вопрос, почему это происходит. Внутри и вблизи материализованных объектов существует информационное поле, имеющее информационно-сотовую структуру. Поток субинформационов, истекающий от каждой структуры, называемой массой, по сути дела образует вокруг массы информационный нимб или информационные следы. Информационное поле, как форма материализованной и дематериализованной информации проникает во все материальные структуры, устанавливая и создавая отношения связи с другими структурами. Находясь в пределах физики, проблему гравитационного поля не удалось решить ни Ньютону, ни Эйнштейну, ни более поздним физикам. Оказалось, что ни гравитон, ни гравитационные волны не являются причиной силы гравитации, поскольку это довольно-таки упрощенные физические конструкции с ограниченными возможностями проявления, которые не подтвердились в реальных экспериментах. В пределах информациологии, используя понятие субинформациалов и информациалов, опираясь на представления теории информационного поля, можно дать исчерпывающее объяснение сущности Всемирного закона тяготения.

Пусть имеются две массы m_1 и m_2 на расстоянии R, как точечные носители информации. Информационно-полевое отношение между этими корреляционными информационными носителями, согласно теореме Гаусса для точечных информационных полей будет иметь вид $\Delta\Phi=4\pi Gm$. Это есть поток информационного поля от точечного носителя m_1, или другими словами — это есть полевой информациал. На расстоянии R плотность информационного потока будет распределяться по сфере $4\pi R^2$ и составит

$$\rho_1 = \frac{\Delta\Phi}{\Delta S} = \frac{4\pi G m_1}{4\pi R^2} = \frac{G m_1}{R^2}.$$

Этот поленой информациал будет воздействовать на массу m_2. Корреляционная связь информационного носителя определяется tm, где t— информационная ковариантa. Тогда получим отношение

$$1 = \rho_1 m_2 = \frac{iGm_1 m_2}{R^2}.$$

Выполнение этого закона для макроструктур материального мира теорией Ньютона предполагалось с мгновенной скоростью. По Ньютону Вселенная имеет абсолютную единовременность. Эйнштейн выступил против этого положения, указав, что всякое воздействие передается со скоростью не превышающей 3×10^8м/сек, но проявление закона инертности масс, обусловленного принципом Э. Маха свидетельствует о мновенной реакции силы инертности на наше возбуждающее воздействие. Т.е., реакция от бесконечно удаленных масс пространства Вселенной следует мгновенно. По Эйнштейну - до них возбуждение дошло бы только через миллиарды лет. Что говорило бы только об некомпактностн, бессвязности Вселенной. Опыты

Н.Л.Козырева и В В.Насонова, проведенные осенью 1978г., а также проверка этих опытов группой академика М. М. Лаврентьева, проведенная на пятидесятидюймовом рефлекторе Крымской астрофизической обсерватории в октябре 1989г. показала, что в реальном мире существует некоторая субстанция которая передается от звёзд практически мгновенно. Группа академика Лаврентьева исследовала дистанционное воздействие от звезд α Lyr на расстостоянии 26,5 световых лет, β Peg — 217,3 световых лет, β And — 75,8 световых лет, δ And — 135,8 световых лет. Опыты показали, что действительно имеется мгновенное дистанционное воздействие этих звезд на детектор. Данная субстанция является причиной силы гравитации и, стало быть, гравитационное возмущение передается мгновенно. С точки зрения информациологии вполне ясно, что такими носителями информации будут идеальные субинформационы. Таким образом, Всемирный закон корреляции имеет мгновенное выполнение. Информационы, как нейтрино, пронизывают все материализованные и дематериализованные участки пространства Вселенной и являются причиной всех мгновенно-действий. Взаимодействие информационных структур в информационном поле практически мгновенное. А поэтому, можно сформулировать следующий закон:

9. Закон генерализационно-единого информационно-сотового взаимодействия.

Взаимодействия материализованных и дематериализованных, виртуальных и гипотетических объектов, тел, частиц, полей и их следов, физических и абсолютных вакуумов, инерциальных и неинерциальных систем выражается Всемирным законом генерализационно — единого информационно-сотового взаимодействия.

$$I = \frac{\xi_{ij} X_i X_j}{R_{ij}^2}$$

Правда, для общей полноты следует сказать, что взаимодействие возможно в поле информационала $4\pi GX_i$, но, поскольку для электродинамических полей $GX_i=e$, то по сути дела, существует два полевых информационала $4\pi GM$ и $4\pi e$. Это факт теоремы Гаусса, который, как показано раньше, прямо приводит к закону Всемирного тяготения, и закону Кулона. Более того, он является исходной позицией Максвелловских законов электродинамики.

Мгновенное выполнение отношений информационного взаимодействия, показывает, что информационное поле Вселенной имеет единый Вселенский детерминизм. Это значит, что каждая информационно-локальная область, каждый составной информациал, как информационная структура, мгновенно определен для всей Вселенной. Следует также учесть, что энергодинамические процессы, такие как распространение электромагнитных волн, фотонных потоков зависят от гравитационной плотности пространства или от гравитационного энергетического информациала.

10. Закон резонансно-волновых воздействий.

При всех резонансно-волновых воздействиях осуществляется дематериализованная передача энерговозмущения.

В пределах материальных структур происходит передача энергии

$$E = kT \sim mc^2 \sim h\nu$$

Для материальных структур передача волны будет проходить с ограниченной скоростью 3×10^8 м/сск поскольку это обусловлено гравитационной плотностью пространства, или его гравитационной вязкостью.

11. *Закон вихревых возмущений*

Вихревые движения информационных полей обуславливают потоки информации вдоль оси вращения.

Закон имеет широчайшее проявление. На макроуровне — это электромагнитные поля, которые описываются уравнениями электродинамики Максвелла

$$\begin{cases} c\,\mathrm{rot}\overline{H} = 4\pi j \\ c\,\mathrm{rot}\overline{E} = -\dot{\overline{B}} \end{cases} \qquad \begin{cases} \mathrm{div}\overline{D} = 4\pi\rho \\ \mathrm{div}\overline{B} = 0 \end{cases}$$

На микроуровне это спин электрона и торсионные поля. ***Существуют как массовые, так и безмассойые торсионные поля.*** Первичными по их обуславливающей причинно-следственной роли являются безмассовые - ***информационные торсионные поля***, *выступающие как совокупность n-мерных спинов различной информационной структуризации.* В последнее время предполагается, что скорость передачи информации посредством торсионных полей может расти до бесконечности, а скорость передачи торсионного возмущения или торсионной ориентации равна бесконечности. Торсионные взаимодействия лежат в основе формирования структурных взаимосвязей и отношений на материальных и нематериальных уровнях. По своей информациологической сути торсионные поля - это информационные поля и они подчинены всем основным законам информациологии: обеспечивается закон сохранения информации, закон равновесия и т.д. Информационное поле посредством различных торсионных полей /спинно-магнитных моментов/ влияют на возникновение материи из вакуума и относятся к критериям материализации и дематериализации.

<u>12. Закон существования белых и черных информационных дыр.</u>

Если в данной локальной области а содержится информация очень высокой плотности, то при разрушении внешним энергетическим полем E=T ее структурных связей произойдет взрывной процесс выхода информации.

Для материализованных структур действие этого закона проявляется при выходе гравитационной массы из коллапса (вспышка сверхновой), при ядерном, термоядерном и обыкновенных взрывах

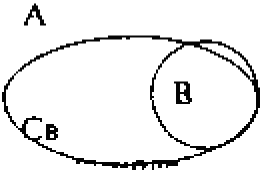
<u>13.Обратная трактовка.</u>

При понижении энергетического информационала E=kT происходит уплотнение информации, ее резкое укомплектование в глубину, провал информационной области в структурные связи.

Примером может служить провал гравитационной массы в коллапс.

<u>14.Закон комплементарности информации</u>.

Ветвящиеся или разделяющиеся информационные структуры делятся согласно принципа дополнительности.



$B \cup C_B = A$

$A_i = \cup (B_i; C_{Bi})$

$I = \sum (I_k - S_k) + \sum S_k$

Закон комплементарности впервые был открыт Нильсом Бором в 1920 году и применен для определения фотона, ках волны и частицы. Этот закон имеет проявление в природе: так, если штабовый постоянный магнит разрезать посредине, то южный конец останется южным полюсом, а на изломе этого куска появится северный полюс. Северный конец останется северным полюсом, а на изломе этого куска появится южный полюс. Деление двойной структуры информационного кода молекулы ДНК обеспечивается этим законом А-Т→А-Т-А-Т. Раздвоение происходит по принципу дополнительности или комплементарности. При обработке числовой информации на счетно-решающих устройствах также не могут обойтись без дополнения к числу: 10 000 равно 9999 + 1111, где 1111 - дополнение к 9999. Вероятностные и корреляционные процессы также основаны на этом законе. Так, если $P_{(a)} = \frac{1}{3}$ — вероятность успеха, то $1 - \frac{1}{3} = \frac{2}{3}$ — вероятность неудачи. Закон комплементарности информации следует из закона сохранения информации.

15. *Обратная трактовка.*

Все информационные поля, массивы, структуры можно сочетать, соединять и соподчинять только согласно закону комплемептариости.

Этот факт очень важен для создания методов генной инженерии, создания идеальных информационных образований при процессах мышления, построении всеобщих нематериальных и материальных информационных структур.

16. Закон сотовости пространства.

Все информационное пространство состоит из бесконечного числа ячеек и сот.

Сотовость — это ячеистость, доменность. Электрон, нейтрон, протон это уже ячейка, сота. Атом, молекула - другая, более крупная и структурирования разновидность соты. Доменные области и кристаллические решетки твердых тел — наиболее наглядное представление сотовости информационных полей. На макроуровне: Солнечная система - сота. Галактика -сота, и, как показывают экспериментальные наблюдения доктора Я. Эйнасто из Эстонии - галактики и их скопления расположены в порядке, который напоминает пчелиные соты огромных размеров. Чем ближе к чашечкам, тем сильнее сконцентрировано вещество.

17. Закон_ кодовости информации

Все информационные структурные образования определены в виде информационных кодов разной степени сложности.

Структуризацию информационных областей под воздействием энергетического информациала F=kT, удобнее всего рассмотреть на примере рекомбинации материи. Согласно существующей космологической теории, после Большого Взрыва происходит адиабатическое расширение области, содержащей вещество. От распада тяжелых X и Y бозонов образуются кварки и антикварки. Во время 10^{-6} до 10^{-5} барионы (протоны, нейтроны и мезоны глюоны, каоны и т.п.). В дальнейшем процессе адиабатического расширения, в результате которого температура падает происходит понижение энергетического информациала - явно выделяются четыре эры.

1 - адронная эра.

При Т $=10^{21}$ К н $\rho=10^{14}$г/см^3. Важной особенностью этой эры есть сосуществование частиц и античастиц. Возникает зарядная асимметрия - разность в количестве нуклонов и антинуклунов составляет 10^{-9} от количества фотонов.

2 — лептонная эра.

Определяется при Т$=10^{12}$ К до 10^9К, при плотности от 10^{14} г/см^3 до 10^4 г/см^3. В начале этого периода энергия равномерно распределена между фотонами, электронами, позитронами, мюонами нейтрино и антинейтрино, которые возникли при аннигиляции электронов и позитронов, согласно схеме e^-+e^+ $v+v$, v — нейтрино. Реакции взаимного превращения протонов в нейтроны происходят по схеме:

$$p+e^- \Leftrightarrow n+v ; p+v \Leftrightarrow n+e^+$$

Но при уменьшении температуры стали более эффективными процессы объединения протонов с нейтронами. В результате этого образуется дейтерий ^2H, тритий ^3H и изотопы ^3He и ^4He. Именно в это время образовалась основная часть гелия, которого сейчас около 30%, а водород составляет около 70 % от всего вещества Вселенной. Это по сути дела была эпоха нуклеосинтеза, т.е. образование гелия происходило во всем объеме пространства Вселенной.

3 - эра излучения или эра фотонной плазмы.

При Т от 10^9 до 3000 К, плотность вещества упала от 1 до 10^{-21} г/см^3. В эту эру температура излучения оставалась такой же, как и температура вещества. Но как только температура достигла значения 3000 К энергия квантов уменьшилась на столько, что они потеряли способность ионизировать атомы водорода. Поэтому процессы рекомбинации электронов с. протонами уже не уравновешивались обратными процессами ионизации. С этого момента главную роль во Вселенной стало играть вещество.

4 — эра вещества.

Началась при указанных условиях и длится до сих пор. Приведенная общеизвестная физическая схема с точки зрения информацилогии явным образом подтверждает действие закона структуризации. Уровень энергетического ирформациала E=kT падает — структуризация возрастает. Действие этого закона можно наблюдать в обыденных условиях. Если в некоторой локальной области происходит понижение энергетического информациала, то структуризация информации в данной области повышается. Ясно видно действие этого закона на таком примере: весной, в течении одних суток, ночью температура падает — вода замерзает, утром температура повышается — лед размерзается, вечером температура понижается — вода снова замерзает. Так втечение одних суток происходят причинно-следственные процессы структуризации и деструктуризации вещества.

Первоосновой природы является фундаментальный микро- и макромерный автоииформгенезис. В окружающем информационном поле постоянно происходят процессы создания и разрушения информационных структур. Фактически это значит, что происходят процессы кодирования и декодирования информации на основе

электромагнитно-резонансного равновесия информационно-сотовых структур.
Рассмотрим автоинформацонный генезис простых структур информациалов, сложных структур-составных информациалов, критерии их становления и разрушения, генетическую память на неорганическом уровне, органическом и создание генетических кодов в процессах живой материи.

Так, при процессе материализации и дематериализации происходит на информационном уровне процесс создания и разрушения кодовых структур. Для неорганического вещества при изменении окружающей температуры возможны изменения агрегатных состояний вещества: оно может быть в состоянии плазмы, пара, твердом состоянии. Все эти состояния обусловлены с одной стороны состоянием энергетико-полевого информациала, с другой стороны внутренним структурным кодом данного вещества. Можно говорить о генетическом коде превращений. Такие простейшие коды свойственны всем элементам таблицы Менделеева, всем неорганическим и органическим соединениям.

Если взять любой химический элемент таблицы Менделеева, то его свойства прежде всего зависят от его информационной структуры. Периодичность, повторение химических свойств, например, зависит от закономерности построения элементарных конфигураций атомов и характеризуется квантовыми числами: главным и орбитальным. Принцип Паули как раз и расписывает электронную структуру каждого химического элемента. По сути с некоторым приближением можно сказать,что электронные формулы Паули выражают структурные коды химических элементов. Для каждого элемента существует также и генетический код его состояний.

Генетический код высокомолекулярных соединений становится безусловно более структурированным и сложным. В результате чего некоторые макромолекулы содержат целые цепи и звенья самоорганизации. Отдельные макромолекулы могут образовывать простейшие ультраструктуры: мембраны, микротрубочки, филоменты (Поглазов, 1966 г). На основе самоструктуризации происходит построение только очень простых компонентов живой системы не способных к самостоятельному существованию. Исследования ряда ученых - Кинга и Вуда, показали, что сборка такой живой системы как фаг из отдельных его компонентов происходит только под контролем генов. Она полностью обусловлена генетическим кодом этой структуры. Воспроизведение себе подобных наблюдается в жидких кристаллах (Евреинова, Опарин 1966). Объект живой жизни значительно сложнее, как структурированный информационный процесс обладает следующими свойствами:

1. Самоорганизации - организация окружающего информационного поля для поддержания устойчивости своей структуры.
2. Самовоспроизводство - порождение подобной себе структуры.
3. Самосовершенствование - всемерное приспособление и адекватизация к внешним изменяющимся информационно-полевым условиям.

Живая система - это органическая система, существование и функционирование которой основано на использовании внутреннего информационного кода. Этот код представляет собой саморазвивающийся информационный процесс, который задает все состояния процесса жизни и развития живого существа. А.Вейсман установил, что носителями генетического кода являются хромосомы. Роль хранителя наследственной информации у всех клеток растений и животных являются ДНК - спирально закрученные одна вокруг другой нити, созданные чередованием четырех типов нуклиотидов и одноцепочная РНК, также состоящая из чередования четырех типов нуклеотидов. Деление клетки как структуры по сути дела обуславливается делением молекулы ДНК, которая происходит на основании информационного закона комплементарности согласно схемы

```
                       А – Ц – Т – Т
                       |   |   |   |
А – Ц – Т – Т          Т – Г – А – А
|   |   |   |    ⇒     А – Ц – Т – Т
Т – Г – А – А          |   |   |   |
                       Т – Г – А – А
```

Таким образом синтез клетки обусловлен тем, что генетический код содержит способность к раздвоению самого кода и передачу этого кода от материнской к дочерней. В этом случае мы приходим к тому, что один структурированный информационный процесс посредством принципа комплементарности порождает другой или множество таких же или подобных структурированных информационных процессов. Зная внутреннюю информационную структуру генетического кода каждого живого существа, каждого живого объекта можно заложить в его состав ядра любой аналогичной яйцеклетки и получить биологический объект с абсолютно идентичными свойствами. Этот этап уже практически достигнут американскими и английскими учеными и он называется клонированием. Следует сказать, что нужно уделить внимание также мутагенным проявлениям.

Они могут быть у живых объектов только тогда, когда в результате внешнего воздействия будет происходить деформация генетического кода потомков. Так, при воздействии радиации, жесткого космического излучения, действия некоторых ядовитых веществ происходит деформация генетических кодов — их мутация. Работы Охагрина и доктора Цзяна рассматривают возможность накладывания определенной информации на сверхвысокие частоты и воздействие этих частот на информационные коды.

Автор новой науки информациологии рассматривая пять основных взаимодействий (гравитационное, електромагнитное, слабое, сильное и цветовое – квантово - хромодинамическое), приходит к выводу о их частном характере и трактует их как различные формы единого первичного (нульсингулярного) взаимодействия, взаимоотношения в природе объединяемого глобальной теорией информационного единства мира. Это обобщающее взаимодействие названо информационно-сотовым со следующими метрологическими характеристиками:

— порядок усиления взаимодействия: 6 (для предыдущих пяти он равен соответственно 1; 2; 3;4;5;);

— границы действия, для слабого см: (соответственно для гравитационного: 10^{-13} - 10^{13} ; для слабого: 10^{-15} и т.д.);

— направление усиления взаимодействий в микромире: 10^{-100} (для гравитационного: 10^{50} , для слабого: 10^{1} и т.д.)

— направление ослабления взаимодействий в микромире: 10^{-100} (для гравитационного 10^{50} , для слабого: 10^{10} и т.д.)

— приближенная (относительная) интенсивность взаимодействия: 10^{100} (для гравитационного: 10^{-40} ; для слабого: 10^{10} и т.д.)

— квант (частица) поля взаимодействий; информацион (для гравитационного: лептон; для сильного: адрон, мезон; для цветкового: кварк, странность, глюон и т.д.);

— частота обмена между частицами: $f_n > f_y$ (для гравитационного: $f_r < f_э$; для слабого: $f_{сл} < f_c$ и т.д.);

— время жизни (полураспада) частиц, с : стабильно (для гравитационного: $>10^{-25}$; для слабого: 10^{-10} и т.д.)

Формула $E=mc^2=hv$ устанавливает связь между материальным объектом массы m и квантом действия h. Фактически так выражается переход от материальной структуры к энергетическому полю, волновому излучению. Согласно принципа неопределенности Гейзенберга в микромире часть привычных в классической физике параметров теряется: они становятся несущественными. Так, в микромире нельзя одновременно говорить о скорости электрона и его координатах в один и тот же момент времени. Классические методы функционального отражения законов в микромире уступают место ***событийным-вероятностным***. Соотношение неопределенностей $\Delta A \times \Delta B > h$, состоит в том, что произведение неопределенностей может быть больше или равно кванту действия, но никогда не становится меньше него. Этим соотношением фактически выделили в физике крайний предел достижимости физики в микромире. ***Все, что по размерам энергодействия меньше оказывается для физики неопределенным.*** Именно поэтому все информационные структуры более глубоких уровней не принадлежат к физически определенному миру, хотя тем не менее имеют объективную реальность. Поэтому в области физики является странной и недоказуемой теория о том, как из безмассовых частиц лентонов, кварков u,d,s,c,b образуются полиморфные информационные структуры, которые являются массовыми частицами (адронами, мезонами, барионами). Может быть в этих отношениях кроются квантовые свойства вещества. Но квантовая теория супергравитации разработанная Д. А. Фридманом, П.Ван Нивехайзеном, С.Фаррарой, Дезером и Зумино все-таки не уяснила причинности гравитации, а также тот факт, как образуется все многообразие реально существующих частиц. В физике появилось целое множество понятий и представлений, которые не представляют собой ни материю, ни энергию - такие как фотоны, пи мезоны, тахионы, нейтрино и т.д. Но все эти понятия - это всего лишь дематериализованная информация, которая в определенных сочетаниях, структурах может быть первоисточником чисто физических понятий энергии, движения, массы. ***Согласно закона информационного равновесия материальные субстанции не появляются ниоткуда и никуда не исчезают***. ***Стабилизация информационного равновесия поддерживается информационно-сотовыми (торсионно-аксиальными) полями*** в которых происходит поляризация информацмонов в направлении создания ***поля кручения*** и поляризация атомов по аксиальному ускорению информационов, что обеспечивает осевое вращение электронных частиц. Любые модели: модель симметрии суперструн и глобальных торсионных полей - это всего лишь ***субмодели информационных структурных образований и информационных полей***.

На данном этапе быстрого динамического развития, широчайшего применения компьютерных систем закономерен процесс перехода от математики абстрактной, выраженной в соответствующих теоремах, абстрактных отношениях, геометрических формах и методах вычислений к ***математике информационной*** — когда целые разделы математики за счет полного программного представления могут восприниматься, как мгновеннодействующий и мгновенно реализующий себя информационал. Прикладные возможности информационной математики разумеется многократно увеличиваются. С точки зрения определения информационных характеристик и параметров проведена хронологическая классификация математических определений, отношений и структур и создана математическая таблица — ИМТ, которая отражает не только генезис абстрактной математики, но и позволяет осуществлять анализ, синтез и давать прогноз по развитию новых способов, методов и моделей. В то же время информациологическая основа математики говорит о том, что еще не созданы такие многофункциональные алгоритмы, методы и модели, которые позволяли бы информационной математике адекватно отражать мир и использовать их в многоаспектном практическом применении. ИМТ, легко переводима в компьютерное представление, обладает однозначностью, чем устранены ненужные семантико-синтаксические неувязки, тавтологии, философские и казуистические многотолкования и, по сути дела, есть некоторой структурной таблицей и различных математических представлений расположенных в порядке их сложности. И. И. Юзвишину удалось разработать информационно-математическую таблицу (ИМТ), которая не только классифицирует и хронологически систематизирует все основные достижения математики, начиная от изобретения натурального ряда чисел в незапамятные времена, через геометрию Евклида (3 в. до н. э.), Булевую алгебру, модель коллектива вычислителей Евреинова, стохастической формулой информации, но также и послужит дальнейшему развитию, углублению и обоснованию таких наук как информационная физика, химия, биология, генетика и т.д.

Информационно-математическая таблица и таблица информационных технологий должны сыграть в информациологии такую же роль, как таблица Менделеева в химии. Особым вниманием должна пользоваться формула информационного кода жизни человека и Вселенной которая имеет вид

$$I = \int_{-\infty}^{\infty} F\left[x_1, x_2, \ldots x_n, t, u, \frac{\partial u}{\partial x_1}, \ldots \frac{\partial u}{\partial x_n}, Q\Phi\right] dV$$

где x_i - зависимые переменные, характеризующие материализованные и дематериализованные объекты *n* - мерного пространства Вселенной, t - время, $u=f(x_i)$ - функция зависимых переменных $x_i=\varphi(\alpha_i)$ от квазинезависимых α_i, $\frac{\partial u}{\partial x_i}$ - скорости и изменения отдельных свойств x_i,

$Q = \psi(\beta_1, \ldots \beta_n)$ – генерализационная функция автоинформгенезиса всех мировых констант, определяемая функционалом Ф.

<u>Основные кванторы математических отношений</u> — ***объединения, суммы, интегралы, дифференциалы, функционалы, векторы, матрицы, тензоры, логические отношения, роторы, дивергенции и градиенты, корреляционные отношения и пр. с успехом используются в информациологии***. Но информациология еще должна прибегать к средствам словесного описания, что говорит о том, что указанных математических структур еще не достаточно для адекватного и полного отражения информации. Стало быть, еще будут проведены новые и более глубокие классификации математических основ и форм на предмет вычленения тех, которые могут быть использованы в информациологии. Более того, информациология стоит на пути создания своих, более адекватных составных информациалов, охватывающих все области имформациниого поля. Ведь пока еще ***никакие математические функционналы*** не в состоянии охватить таких областей информации отражаемые человеческими чувствами как ***обоняние, вкус, музыка, чувство тепла, холода, голода, страха, голода, боли, ненависти***. Фактически существуют еще многие информационные области, которые непредставичы в мире математики, но представляют собой информационные поля отражаемые относительно разума человека.

Информациология позволяет представить всю Вселенную, как логическую связь информационных отношений выраженную в соответствующей логической блок-схеме книги И.И.Юзвишина "Информациология", рис.2.2. переходных процессов информационной модели Вселенной. Определено, что из ***отношений <u>пространства и времени</u> возникает <u>движение</u>, а отношение <u>материи и скорости</u> представляют <u>энергию</u>. Соотношения <u>энергии, движения, массы и антимассы</u> представляют <u>информацию</u>***. Поскольку информация является основой макро- и микромерных процессов, то ее можно записать в виде всеобщего Вселенского кода, определяемого функционалом

$$I(t) = F[f(Э,Д,М,А),t].$$

Движение - это $V = \frac{L}{t}$,

энергия - это

$$Э = \frac{m}{(V^2)^{-1}}, \quad I=((Э/Д)/М)/А.$$

Логическая информационная модель Вселенной будет иметь вид:

|→ИНФОРМАЦИЯ (ОТНОШЕНИЯ) → СВЯЗИ → ВЗАИМОЗАВИСИМОСТИ →
|СУПЕРАВТОКОРРЕЛЯЦИЯ → ВЗАИМОПРОНИКНОВЕНИЯ → ЭНЕРГИЯ →
|ВЗАИМОДЕЙСТВИЯ → МАТЕРИАЛИЗАЦИЯ →ДЕМАТЕРИАЛИЗАЦИЯ →
|РИТМИЧНОСТЬ → ПЕРИОДИЧНОСТЬ →ВЕЩЕСТВО (МАССА) →
|НТИВЕЩЕСТВО (АНТИМАССА) → МАТЕРИЯ → АНТИМАТЕРИЯ →
|ПРОСТРАНСТВО → ВРЕМЯ → ДВИЖЕНИЕ→|.

В работе "К обоснованию фундаментальных основ информациолегии" /15/ И.И.Юзвишин, опираясь на данные "Энциклопедического словаря" и материалов таблицы астрофизических данных приходит к выводу, что объем всей материи пространства Вселенной составляет 0,000001% определенного пространства Вселенной. В микромире картина аналогичная. Вся масса атома сосредоточена в ядре размеры которого в 10^5 раз меньше атома. Строение атома показывает, что атом в основном состоит из пустоты. Макромиры и микромиры, как материальные структуры, плавают в безночально- бесконечных пространствах информационного поля Вселенной. 99,999999% - определенного пространства таким образом представляет нечто, что с точки зрения физики может быть трактовано как физический вакуум или пустота. Экспериментально обнаружено, что при тех условиях, когда среднее число частиц в вакууме равно нулю, там происходит рождение виртуальных частиц, которые влияют на физические процессы. Более того, это информационное поле является фоном, фундаментом и первоосновой, которая порождает материальные объекты. Таким образом ***информационное поле* — это фундаментальный субстрат мироздания.**

Открытие сущности информации как фундаментального субстрата континуумных элементарных отношений, автоинформгенезиса и новых всеобщих законов информациологии позволяет с информациологической точки зрения провести исследования Вселенной, ее видимых и невидимых законов, определяемых на материальном уровне и дематериализованных критериев, которые обуславливают материализацию информационного поля. Это исследование ставит целью вскрыть причинно-следственные критерии - каким образом из дематериализованной информации, из супервысокого или абсолютного физического вакуума, где по сути дела отсутствуют материальные частицы происходит зарождение материальных частиц? Каким образом нульмассовые и нульзарядные информационные субстанции рождают массовы зарядные частицы протоны, электроны и нейтроны.

Информациологическая модель Вселенной, с учетом вышесказанных положений - это в строгом смысле слова не космологическая модель Вселенной, так как в ее основе не заложены пространственно-геометрические представления модели Фридмана, или иных моделей, а информациологические представления о структуризации информации, которые тем или иным образом реализуют себя на краевых условиях информационных областей. Краевые условия информационных областей наиболее удобно выделить по ***двум параметрам*** - для макромира - по средней плотности вещества представляемого в г/см^3, для микромира - по кванту энергодействия. В работе /15/ приведена закономерно-циклическая система автоинформгенезиса материализации и дематериализации Вселенной, в которой отражен целый спектр краевых условий от абсолютно твердого вещества до супервысокого вакуума. Таблица содержит 25 различных значений плотности вещества и, следовательно разбивает все информационное поле Вселенной на 25 областей, которые через радиус вектор понижения плотности представимы в виде концентрических колец. Для удобства представления выделим только некоторые, наиболее радикальные информационные области, соотвеествующие следующим краевым условиям:

Краевые условия информационных областей

Плотность вещества, г/см^3	Состояние вещества
$10^{20} \leq \rho_1 \leq 10^{40}$	Колаптическое состояние вещества
$10^{20} \leq \rho_2 \leq 10^{6}$	Высокоплотное вещество
$10^{6} \leq \rho_3 \leq 1$	Вещество средней плотности
$1 \leq \rho_4 \leq 5\times10^{-8}$	Слабый лабораторный вакуум
$10^{-18} \leq \rho_5 \leq 10^{-31}$ $h=10^{-34}$	Граница кванта энергодействия
$\rho_6 \leq 10^{-40}$	Сверхвысокий вакуум

В пределах этих краевых условии информационных областей материальная структура определяет себя различной степенью структурированости: просто как элементарная частица, атом, молекулы газа, жидкость, твердое вещество, вещество с плотностью нейтронной звезды, вещество с плотностью черной дыры и сингулярное состояние вещества. Чем выше степень структуризации, тем, вероятнее всего, выше плотность информации данной структуры. Совершенно понятно, что один кубический сантиметр вещества нейтронной звезды будет иметь большую плотность информации, чем один кубический сантиметр физического вакуума. В этом случае очивидна следующая зависимость $i\rho_i\alpha = \rho$ - где $\rho_i\alpha$ - это плотность информации, i - информационная ковариантa, ρ - физическая плотность краевых условий.

<u>Краевые условия информационных областей</u>, это те условия, которые позволяют реализоваться тем или иным информационным структурам, под действием определенных энергетико-полевых критериев. Так, если $10^{20} \leq \rho_1 \leq 10^{40}$, то действуют только энергетико-полевые критерии 4πe, 4πGM. Когда $10^{6} \leq \rho_2 \leq 10^{20}$, то действуют энергетико-полевые критерии 4πe, 4πGM, E=kT, $E=mc^2$, E=hv. Когда $10^{-18} \leq \rho_5 \leq 10^{-31}$, то действуют энергетико-полевые критерии E=kT, E= hv. Если взять значение кванта энергодействия H≤h, то этим будет осуществлен переход в физически неопределенный мир - калибровочную пустыню. С точки зрения информации - переход будет осуществлен в информационное поле, где H следует понимать, как организующий квант информации. Следует обратить внимание на то, что физический вакуум и калибровочная пустыня <u>за пределами кванта энергодействия с информационной точки зрения значит одно и то же.</u> Учитывая изложенное, топологически информационную модель Вселенной можно представить так: в центре поместить структурные образования самой высокой плотности. Для материальных структур, получим четыре концентрических окружности твердого вещества и две концентрические окружности газообразного вещества. Седьмой окружностью будет граница физического мира, отвечающая кванту энергодействня.

Согласно физических представлений все, что меньше кванта энергодействия физически не определяемо. То есть, за седьмой окружностью будет находиться огромное информационное поле дематериализованной информации.

Все информационные поля бесконечной Вселенной разбиваются на материализованную и дематериализованную область. Материализованная область имеет ограниченные размеры, дематериализованная - беспредельна. Положения информациологин позволяют рассматривать информационные структуры, которые находятся за пределами кванта энергодействия в поле действия законов информациологин. Согласно этому в информациологии можно говорить о нульмассовых и нульзарядных состояниях вакуума. Если в физическом вакууме существует обьем вакуума $V_в$, то правомерно допустить, что этот объем представляет пространство, в котором всегда найдется какая-либо точка этого пространства, начало координат какой-то системы в текущем времени. Таким образом, при наличии пространственно-временных факторов можно говорить о скоростном градиенте вакуума во времени. Используя функцию Гамильтона в нуль массовых и нульзлрядных полях имеют место конструкции $\frac{\partial}{\partial t};\frac{\partial^2}{\partial t^2}$. От центра информационной области проведем информационный, градиент $\frac{\partial \rho}{\partial t}$, который на физической границе удовлетворяет следующим краевым условиям:

$$1).\ \frac{\partial \rho}{\partial t}\Big|_{H<h} = m_0;\ \frac{\partial \rho}{\partial t}\Big|_{H\le h} = q_0 \qquad 2).\ \frac{\partial |\rho_i \alpha|}{\partial t}\Big|_{H\le h} = |\rho_i \alpha|,$$

которые обозначают, что вещественные субстанции переходят в информационные поля Первые сотношения говорят о переходе вещества к нульмассе и нульзаряду, а второе говорит о том, что при переходе физической границы информационные поля остаются. Понятие нульмассовости и нульзарядности не следует понимать как нульмассовость и нульзарядность пустого множества, так как может оказаться, что это равновесная нуль масса и равновесный нульзаряд. Аналогия этому очень простая: следует взять нуль как начало координат числовой оси.

Такой нуль совершенно не значит, что на числовой оси нет ни одного числа, что множество чисел - пустое множество Ø. Такой нуль всегда можно представить как: 4-4=0 или 100-100=0, и т.п. В таком случае нульмасса и нульзаряд будут иметь выражение через массу – антимассу

$$m_0 = \sum_{i=1}^{\infty} m_i - \sum_{i=1}^{\infty} m_i$$

заряд-антизаряд

$$q_0 = \sum_{i=1}^{\infty} q_i - \sum_{i=1}^{\infty} q_i$$

Если мы выделим некоторый обьем, принадлежащий дематериализованному информационному полю, то в декартовой системе координат, когда $\Delta l^3 = \Delta x \times \Delta y \times \Delta z$, будем иметь:

$$\frac{\partial l}{\partial t}$$

- скорость,

$$m_0 \frac{\partial l}{\partial t}$$

- импульс,

$$\frac{\partial^2 l}{\partial t} m_0 ; \frac{\partial l}{\partial t} q_0$$

- некоторая информационная структура, носящая в себе массовое и зарядное квазиэнергетическое начало.

Фактически можно предполагать, что существует какие-то корреляционные отношения, подобные отношениям между выражениями

$$\bigcup_{n=1}^{\infty} m_{0n} \frac{\partial^2 l}{\partial t} \approx i H_i v_i ,$$

где i - некоторая информационная коварианта.

Такого рода выражение говорит о том, что время для таких информационных полей не остановилось, а течет, что обеспечивает выполнение закона об изменяемости информационных полей. При этом ясно подчеркивается тот факт, что дематериализованный информациал $H_i v_i$ - это объединение энергетических информационных субстанций. Вместе с тем, согласно условия $\frac{\partial H_i v_i}{\partial t} = \infty$, отражает факт мгновеннодействия дематериализованных информационных полей. В пределах связных частот: для границы физического мира 10^{-34}, $\nu=\nu_{max}$, для субинформациона 10^{-100}, $\nu=\nu_{min}$. В этом огромнейшем информационном поле, которое распространяется в бесконечность, под воздействием определенных информационных критериев формируются информационные образования и структуры, опираясь на которые на физической границе h происходит материализация. Определим информационную емкость информационных полей. Опираясь на определение И.И. Юзвишиным информациона, можно сказать, что все информационы, как атрибуты i-х и j-х отношений самих с собой или между собой можно записать в виде $aiRai \rightarrow aii$; $bjRbj \rightarrow bjj$; $aiRbj \rightarrow aibj$; В декартовой системе координат расположив по оси ОХ — информационы (0, a_1,a_2,..., a_n), по ОУ - (0,$в_1$,$в_2$,...,$в_n$), по OZ - (0,c_1,c_2,...,c_n) тогда внутри каждого пространственного информационного куба размерности п будут расположены отношения $(0,0,0);\ (в_1c_1, a_1c_1, a_1в_1); \ldots ;(в_nc_n, a_nc_n, a_nв_n)$. По главной диагонали будут расположены элементы $(0,0,0);\ (в_1c_1, a_1c_1, в_1c_1); \ldots ;(в_nc_n, a_nc_n, a_nв_n)$. Общее количество всех информационных позиций в этом декартовом информационном кубе будет составлять $N=n^3$. Общее число всех информационных состояний в этом информационном кубе - это множество подмножеств от общего числа элементов N. Согласно комбинаторики множество подмножеств, как известно, составляет $I=2^N=2^{n3}$. Так, если имеем $n = 3$, (а,в,с), тогда общее число информационных состояний будет $I = 2^{27}$. Фактически это размер данной информационной области. Если предположить, что компактно заполнен электронами с радиусом 10^{-15} м, то информационное поле отношений между электронами будет равно: $I = 2^{10^{(15)3}} = 2^{10^{45}}$. Размер идеального субинформациона равен 10^{-100} . Тогда в одном кубическом метре пространства содержится $N = 10^{300}$ идеальных информационов. Таким образом, кубический метр физического вакуума не пуст, а содержит $I = 2^{10^{300}}$ информационных состоянии объектов типа $a_i; \frac{\partial a_i}{\partial t}; m_0 \frac{\partial a_i}{\partial t}; m_0 \frac{\partial^2 a_i}{\partial t^2}; iH, v.$ Если обьем определенного пространства обозначить $V=100\%$, а обьем холодного пространства (сверхвысокий космический вакуум) $V_{хп}$=99,999999999%, то получим следующую формулу информационной насыщенности физически ваккумного пространства Вселенной $V = V_{хп} \times 2^{10^{300}} = 99{,}99999999 \times 2^{10^{300}}\%$. То, что в физике воспринимается как нустота, как физический вакуум с информационной точки зрения представляет собой чрезвычайно насыщенное информационное поле в котором только отношений между идеальными субинформационами будет $V = 99{,}999999999 \times 2^{10^{300}}\%$. Такое поле и обуславливает появление будущих провозвестников кирпичиков изученной материи. Материализацию информационных полей можно представить как результат воздействия кванта энергодействия h на нуль массу или нуль заряд. Следует признать следующие равенства: m=moh; $q=q_0h$; $h(H,v_i)=hv$. Тогда получим материализованные отношения

$$a_i; \frac{\partial a_i}{\partial t}; m \frac{\partial a_i}{\partial t}; m \frac{\partial^2 a_i}{\partial t^2}; hv.$$

Обьем холодного пространства Вселенной значительно больше теплового, причем $V_{хп}=10^{-90}/_0V_в$. Обьем вещества во Вселенной

$V_вB=10^{-17}V_в$.

Основной вопрос философии о первичности материального или идеального в том понимании, что материя - субстрат только вещественного мира решается в пользу объективного идеализма, который утверждает наличие духовного невещественного, нематериального первоначала природы вне зависимо от сознания. Субъективный идеализм состоит из информациализма в следующем отношении: ***существует объективное информационное поле за пределами физического мира и существует <u>субъективное информационное поле в голове человека</u>, как <u>отражение познанных материализованных и дематериализованных информационных полей</u>.*** Это отражение представляет собой знание о материализованном и дематериализованной мире. Поскольку, как было доказано раньше, эти поля незамкнуты, то это по сути дела значит, что позиции агностицизма необоснованы. Информациология доказывает данное положения просто и ясно, опираясь на фундаментальную аргументацию.

Информациология, используя достижения физики, химии, математики, механики, астрофизики, биологии, генетики, космологии, истории — совокупность всех фундаментальных и социально-исторических наук, в тоже время развивает свои собственные методы изучения отношений между свойствами, качествами, характеристиками, особенностями признаками и явлениями материализованных и дематериализованных объектов, их элементов, полей, следов, возбуждений, оцеляций и импульсов, создает новые общие структурные формы и виды информациологических отношений и законов, которые экстраполируется на краевых условиях информационных областей, как конкретные законы, формулы, явления общеизвестных наук. Методами информациологии удалось выявить информациологические противоречия законов Ньютона, теории относительности Эйнштейна, кибернетики Виннера, показать конструктивную роль систематизации периодичности таблицы Менделеева, показать глубинные основы явления периодичности в целом. В рамках информационной модели введено понятие ***<u>информациона</u>, под которым понимается бесструктурный, элементарный квант отношений микро и макродинамических процессов и явлений — это <u>субэлементарная информационная частица</u>.***

Среди научных и практических направлений, которые включают в себя информациологию (их около сорока и перечень их будет расщряться), можно отметить такие важные, как определение ***информационной единицы измерения и информационной системы координат для энергии, движения массы, пространства и времени***;

разработка унифицированных единиц измерения информации в физике, химии, биологии, астрономии, космологии, технике, связи. медицине, социальной сфере, в сельском хозяйстве, системах автоматизированного и автоматического управления, компьютерно-сетевой технологии.

Рассмотрим, предложенную А.В. Киндеревичем, унификацию единиц измерения физических величин. Согласно справочника "Физические величины" под редакцией И.С.Григорьева, Е.З.Мейлихова 1991г., Москва, энергоатомиздат, единицы физических величин, допущенные к применению ... их наименования и обозначения установлены государственным стандартом "ГСИ: Единица физических Величин" ГОСТ 8.4I7-81 (СТ СЭВ 1052-78). Основные величины и единицы СИ.

1. Длина *l*, единица - метр.
2 Масса m, единица - килограмм.
3. Время t, единица - секунда.
4. Сила электрического тока I, единица - ампер.
5. Термодинамическая температура Т, единица - Кельвин.
6. Количество вещества n, единица - моль.
7. Сила света I, единица - кандела.

Кроме основных единиц существует дополнительные единицы:
Плоский угол - α, единица - радиан.
Телесный угол Ω, единица – стерадиан,
и целая совокупность производных единиц системы СИ.

Попытаемся все основные физические величины выразить через информационное понятие длины, т.е. через метр. Поскольку, все производные единицы выражаются через основные, то, по сути дела, получим факт приведения всех единиц СИ к информационному понятию длины. Такого рода унификация послужит первым шагом к унификации всех единиц измерения и всех констант материального мира к некоторому информационному эталоно-кванту.

Метр представляет собой расстояние проходимое в вакууме плоской электромагнитной волной за 1/299792458 долю секунды. Следует сказать, что в астрономии с чисто практических мотивов, астрономическая единица, световой год и парсек уже приведены к метрическому выражению.

1 а.е. $=1{,}49\times10^{11}$ м
1 св.год $= 9{,}46\times10^{15}$ м
1 парсек $= 3{,}0\times10^{16}$ м.

Время

Согласно уравнения Г. Минковского

$$x^2+y^2+z^2=c^2t^2,\ \text{где}\ c^2t^2=R^2 \Rightarrow ct=R\ \text{получим}\ 1\,\text{сек}=\frac{1}{3\times10^8}\,м=3{,}3\times10^{-9}\,м.$$

Масса

Используя известные формулы $E=mc^2=h\nu$, $m=\frac{h\nu}{c^2}$, полагая при этом, что $\nu=3{,}0\times10^{14}$ - частота зеленого спектра получим:

$$1кг=\frac{6{,}62\times10^{-34}}{9\times10^{16}}\times3\times10^{14}\times3\times10^{8}\,m=6{,}62\times10^{-40}\,m$$

В данном случае было использовано то, что $1Гц=\frac{1}{c}$, а так же основное свойство информациологии - реализация информации на конкретных краевых условиях.

Количество вещества.

Из молекулярно-кинетической теории известно, что количество вещества $n=\frac{N}{V}$, где N - число молекул, V - объем. Из формулы p=nkT имеем $n=\frac{p}{kT}$. Подставив следующие конкретные значения краевых условий $p_0=1{,}013\times10^5$ Па; $T_0=273$К; $k=1{,}38\times10^{-23}$ Дж/К, получим:

$$n=\frac{1{,}01\times10^5}{1{,}38\times10^{-23}\times273}=2{,}69\times10^{25}\,м^{-3}.\ 1моль=2{,}69\times10^{25}\,м^{-3}.$$

Температура.

Единицей измерения температуры будет 1 кельвин. Воспользуемся приведенным выше сотношением из которого $273К=\frac{np}{k}$ или

$$1кельвин=\frac{np}{k273}=\frac{2{,}69\times10^{-25}\times1{,}01\times10^5}{1{,}38\times10^{-23}\times273}=7{,}9\left(м^{-3}\right)^2=7{,}9м^{-6}$$

1 кельвин=7,9 $м^{-6}$

Сила тока.

Из соотношений Q=It, Q=cmT, при краевых условиях удельной теплоемкости воды будем иметь

$$I = \frac{cmT}{t}, \ 1 \text{ ампер} =$$

$$= \frac{4200 \times 6{,}6 \times 10^{-40} м}{\dfrac{1}{3 \times 10^{8} м}} = 4200 \times 6{,}6 \times 10^{-40} \times 3 \times 10^{8} м^{2} = 8{,}3 \times 10^{-28} м^{2}$$

<u>Сила света.</u>

Сила излучения, согласно справочника стр. 15,

$$I_e = \frac{d\Phi_e}{d\Omega}, \text{ но } \Phi e = \frac{dW}{dt}, \text{ тогда } I_e = \frac{d^2W}{dtd\Omega}$$

Кандела равна силе света в заданном направлении источника испускающего монохроматическое излучение частотой 540×10^{12} Гц, сила излучения которого в этом направлении составляет 1/683 Вт/ср. Телесный угол Ω, как пространственная геометрическая форма выражается через линейные метры.

$$\frac{R}{\ell} = \sin\Omega; \ \Omega = \arcsin\frac{R}{\ell}.$$

Так, если R = l, l = 2 , то $\Omega = \frac{\pi}{6}$.

Выражение

$$\frac{d^2W}{dtd\Omega} \sim \frac{h\nu 6}{1\pi};$$

$$d^2W/dt\, d\Omega \sim W/t\, \Omega \sim h\nu\, 6 / 1 \text{ сек } \pi;$$

1 кандела = ($6{,}62\times10^{-34}$ Дж с $5{,}4\times10^{14}$ Гц 6)/ 1 сек π ;

Выразим Дж в метрическом представлении A=FS.

$$1 \text{ Дж} = 1 \text{ кг } 1 \text{ м} / 1 с^2 \text{ м} = 1 \text{кг м}^2/с^2 =$$
$$= 6{,}62\ 10^{-40} м^3\ 9\ 10^{16} м^2 = 3{,}1\ 10^{-23} м^5;$$

Тогда

$$1 \text{ кандела} = (6{,}62\ 10^{-34}\ 5{,}4\ 10^{14}\ 3\ 10^{8}\ 6\ 3{,}1\ 10^{-23} м^5) / 3{,}14 =$$
$$6{,}9\ 10^{-11} м^5.$$

Вывод: Все основные измерения физических величин приводятся к информационной единице – метру. Правда, метр будет фигурировать в разных степенях: в минус третьей, шестой, и т.д. Приведение производных единиц физических величин даст также и дробные степени. Все это говорит только о том, что представляя некоторый физический закон или явление, информационная единица - метр учавствует в очень сложных отношениях структурированности.

Список литературы

1. Ньютон И. Собрание трудов. M-JI. 1936г.
2. Эйнштейн Л. Собрание научных трудов. М. 1965г.
3. Ландау Л., Лифшиц Е. Квантовая механика. М.1989г.
4. Зельдович Я., Холопов М. Драма идей в познании природы.
5. Евреинов Э.В. Однородные вычислительные системы
6. Евреинов Э.В., Юзвишин И.И. Единая распределённая информационно-сотовая сеть финансово-комерчсских банков мира. Журнал НТБ №1 МАИ 1995г.
7. Паули В. Теория относительности. М. 1983г.
8. Пригожий И., Стенгерс И. Время, хаос, квант. М. 1994г.
9. Юзвишин ИИ. Информациология. М. 1993г.
10. Юзвишин ИИ. Информациология. Уч. пособие. М. 1994г.
11. Козырев НА Вспыхивающие звёзды. Ереван. 1977г.
12. Козырев Н.А., Насонов В.В. Астрономия и небесная механика. М. 1978г.
13. Лаврентьев М.М., Еганова И.А. О дистанционном воздействии звезд на резистор. 1989г.
14. Киндеревич А.В. Теория поля. Ее применение для создания сверхскоростных космических летательных аппаратов, гравитационных электростанций, осуществления холодного ядерного синтеза. Киев. 1998г.
15. Проблемы информациологии. Москва 1998г.

ЛИТЕРАТУРА

1. *Агудов В. В.* Количество, качество, структура.— «Вопросы философии», 1967, № 1.

2. *Адамар Ж.* Исследование психологии процесса изобретения в области математики. М., 1970.

3. *Адров В. М.* Информация в неживой природе.— «Ученые записки МГПИ им. В. И. Ленина», 1968, № 290.

4. *Акчурин И. А.* Развитие кибернетики п диалектика.— «Вопросы философии», 1965, № 7.

5. *Амосов Н. М.* Искусственный разум. Киев, 1969.

6. *Амосов Н. М.* Моделирование информации и программ в сложных системах.— «Вопросы философии», 1963, № 12.

7. *Анисимов С. Ф.* Состношение категорий закона, причинности, - необходимости, случайности.— «Вопросы философии», 1955, № 6.

8. *Анисимов С. Ф.* Человек и машина. М., 1959.

9. *Анохин П. К.* Опережающее отражение действительности.— «Вопросы философии», 1962, № 7.

10. *Аптер М.* Кибернетика и развитие. М., 1970.

11. *Арманд А. Д.* Природные комплексы как саморегулируемые информационные системы.— «Известия Академии наук СССР. Серия географическая», 1966, № 2.

12. *Аход А. В., Улман Д. Д* - Теория языков.— В кн.: Кибернетический сборник, вып. 6. М., 1969.

13. *Баженов Л. Б.* О некоторых философских аспектах проблемы моделирования мышления кибернетическими устройствами.—
В кн.: Кибернетика, мышление, жизнь. М., 1964.

14. *Баженов Л. Б.* Физика и теория информации.— «Вопросы философии», 1961, № 8.

15. *Барановский Н. В.* О причинной обусловленности случайности.— «Философские науки», 1963, № 6.

13. *Берг А. И.* Кибернетика — наука об оптимальном управлении. М.—Л., 1964.

17. *Беркс А. У.* Вычисление, поведение и структура неизменных и растущих автоматов.— В кн.: Самоорганизующиеся системы. М., 1964.

18. *Бернар К.* Введение к изучению опытной медицины. СПб.— М., 1866.

19. *Берталанфи Л.* Общая теория систем — критический обзор.— В кн.: Исследования по общей теории систем. Ежегодник — 1968. М., 1969.

20. *Бирюков Б. В.* Кибернетика и методология науки. М., 1974.

21. *Бирюков Б. В.* Машина и мышление (три принципа).— В кн.: Художественное и научное творчество. Л., 1972.

22. *Бирюков Б. В., Геллер Е. С.* Кибернетика в гуманитарных науках. М., 1973.

23. *Бирюков Б. В., Геллер Е. С.* Моделирование и нейрокибернетика.— «Природа», 1967, № 4.

24. *Бирюков Б. В., Геллер Е. С.* О кибернетическом моделировании познавательных психических процессов. Предисловие.— В кн.: *Новик И. Б.* Философские вопросы моделирования психики. М., 1969.

25. *Бирюков Б. В., Урсул А. Д.* К проблеме объективности информации.— В кн.: Методологические проблемы кибернетики.

(Материалы к Всесоюзной конференции), т. 1. М., 1970.

26. *Бишоп Дж.* Обратная связь через окружение как аналог функционирования мозга.— В кн.: Самоорганизующиеся системы. М., 1964.

27. *Блауберг И. В., Садовский В. И., Юдин Э. Г.* Системный подход в современной науке.— В кн.: Проблемы методологии системного исследования. М., 1970.

28. *Блохинцев Д. И.* Критика идеалистического понимания квантовой механики.— «Успехи физических наук», т. 45, 1951,

вып. 2.

29. *Богданов А. А.* Всеобщая организационная наука (тектология), т. I, М., 1925; т. II, Берлин, 1927; т. III, Берлин, 1929.

30. *Больцман Л.* Лекции по теории газов. М., 1953.

31. *Бонгард М. М.* О понятии «полезная информация».— В кн.: Проблемы кибернетики. М., 1963, вып. 9.

32. *Борель Э.* Вероятность и достоверность. М., 1961.

33. *Бори М.* Действительно ли классическая механика детерминистична? — В кн.: *Борн М.* Физика в жизни моего поколения.

М., 1963.

34. *Бриллюэн Л.* Наука и теория информации. М., 1960.

35. *Бриллюэн Л.* Научная неопределенность и информация. М., 1966.

36. *Бродский И. Н.* Причинность и информация.— «Вестник ЛГУ. Серия экономики, философии, права», 1963, № 17, вып. 3.

37. *Бройлъ Л.* По тропам науки. М., 1962.

38. *Бэкон Ф.* Новый Органон,— В кн.: *Бэкон Ф.* Сочинения, т. 2. М., 1972.

39. *Вавилов С. И.* Собрание сочинений, т. III. М., 1956.

40. *Винер Н.* Кибернетика и общество, М., 1958.

41. *Винер И.* Кибернетика, или Управление и связь в животном и машине. М., 1968.

42. *Винер И.* Я — математик. М., 1967.

43. *Войшвилло Е. К.* Попытка семантической интерпретации статистических понятий информации и энтропии.— «Научно-техническая информация», 1963, № 10.

44. *Воробьева М. Б.* Об этапах развития французского математического языка.— В кн.: Особенности языка научной литературы. М., 1965.

45. *Гальперин П. Я.* Психология мышления и учение о поэтапном формировании умственных действий.— В кн.: Исследование мышления в советской психологии. М., 1966.

46. *Гейзенберг В.* Философские проблемы атомной физики. М., 1953.

47. *Геодакян В. А.* О структуре эволюционирующих систем.— В кн.: Проблемы кибернетики, вып. 25. М., 1972.

48. *Гиббс Дж. В.* Основные принципы статистической механики. М., 1946.

49. *Гиббс Дж. В.* Термодинамические работы. М., 1960.

50. *Глезер В. Д., Цукерман И., И.* Информация и зрение. М., 1961. ;

51. *Глушков В. М.* Введение в кибернетику. Киев, 1963.

52. *Глушков В. М.* Мышление и кибернетика.— «Вопросы философии», 1963, № 1.

53. *Гнеденко Б. В.* Курс теории вероятностей. М., 1967.

54. *Голдман С.* Теория информации. М., 1957.

55. *Гришкин И. И.* Понятие информации. Логико-методологический аспект. М., 1973.

56. *Гриманин Б. А.* Учет стоимости информации в теории ценности информации.— «Известия АН СССР. Техническая кибернетика», 1967, № 2.

57. *Гутчин И. Б.* Кибернетические модели творчества. М., 1969.

58. *Гущин Д. А.* Категория информации и некоторые проблемы развития.— «Вестник ЛГУ. Серия экономики, философии и права», 1967, № 23, вып. 4.

59. *Дарвин Ч., Уоллес А. Р. О* стремлении видов образовывать разновидности и сохранении разновидностей и видов естественными способами отбора.— В кн.: *Дарвин Ч.* Сочинения, т. 3. М., 1939.

60. *Детловс В.* Статистический анализ гармонии.— В кн.: Точные методы в исследовании культуры и искусства. (Материалы к симпозиуму), т. 2. М., 1971.

61. *Дирак П. А. М.* Принципы квантовой механики. М., I960.

62. *Дубровский Д. И.* Психические явления и мозг. М., 1971.

63. *Дюкрок А.* Физика кибернетики.— В кн.: Кибернетика ожидаемая и кибернетика неожиданная. М., 1968.

64. *Жуков Н. И.* Информация. Философский анализ. Минск, 1968.

65. *Завадский С. А.* Теория и практика «машинкой поэзии». (О некоторых зарубежных экспериментах).— В кн.: Художественное и научное творчество. Л., 1972.

66. *Зарипов Р. X.* Кибернетика и музыка. М., 1971.

67. *Земан И.* Познание и информация. М., 1965.

68. *Иванов В. В.* Бинарные структуры в семиотических системах.—В кн.: Системные исследования. Ежегодник— 1972. М., 1972.

69. *Ильин В. А.* Некоторые вопросы науки о системах управления. В кн.: Философские вопросы кибернетики. М., 1961.

70. Инженерная психология. Под ред. А. Н. Леонтьева. М., 1964.

71. История античной диалектики. М., 1972.

72. *Ичас М.* Биологический код. М., 1971.

73. *Кальвин М.* Химическая эволюция. М., 1971.

74. *Кастлер Г.* Возникновение биологической организации. М.,

75. *Клюейвер А., Ван-Нилъ К.* Вклад микробов в биологию. М., 1959.

76. *Кобозев И. И.* Исследование в области термодинамики процессов, информации и мышления. М., 1971.

77. *Кобозев Н. И.* О физико-химическом моделировании процессов информации и мышления.— «Журнал физической химии», т. 40, 1966, № 2—3.

78. *Колбановский В. Н.* О некоторых спорных вопросах кибернетики.— В кн.: Философские вопросы кибернетики. М., 1981.

79. *Колмогоров А. Н.* Жизнь и мышление как особые формы существования материи.— В кн.: О сущности жизни. М., 1964.

80. *Колмогоров А. Н.* Теория передачи информации. М., 1956.

81. *Кольман Э.* О философских и социальных проблемах кибернетики.— В. кн. Философские вопросы кибернетики. М., 1961.

82. *Котельников В. А., Николаев А. М.* Основы радиотехники. М., 1954.

83. *Кочергин А.Н.* Моделирование мышления. М., 1969.

84. *Кравец А. С.* Вероятность и системы. Воронеж, 1970.

85. *Крамер Г.* Математические методы статистики. М., 1948.

86. *Крылов Н. С.* Работы по обоснованию статистической физики. М.— Л., 1950.

87. *Кузнецов И. В.* Важные проблемы научного познания. Послесловие к книге: *Бриллюзн Л.* Научная неопределенность и информация. М., 1966.

88. *Кулагина О. С, Ляпунов А. А.* К вопросу о моделировании эволюционного процесса.— В кн.: Проблемы кибернетики, вып. 16. М., 1966.

89. *Куликова И. С.* Что такое театр абсурда? М., 1967.

90. *Кэмпбелл Д. Т.* Слепые вариации и селективный отбор как главная стратегия процессов познания.— В кн.: Самоорганизующиеся системы. М., 1964.

91. *Ландау Л. Д., Лифшиц Е. М.* Статистическая физика. М., 1964.

92. *Лаплас П. С.* Опыт философии теории вероятностей. М., 1908.

93. *Лахтин Л. К.* О методе Пирсона в приложении к задачам статистики и биологии. М., 1904.

94. *Лебедев Д. С, Левитин Л. Б.* Перенос информации электромагнитным полем.— В кн.: Проблемы передачи информации, вып. 16. М., 1964.

95. *Левин Б. Р.* Теория случайных процессов и ее применение в радиотехнике. М., 1957.

96. *Левит А.* О противоречии между содержанием и формой теории информации.— «Научная мысль», 1968, № 11.

97. *Лем С.* Сумма технологии. М., 1968.

98. *Леонтьев А. Н., Кринчик Е. Л.* О применении теории информации в конкретно-психологических исследованиях.— «Вопросы психологии», 1961, № 5.

99. *Лотман Ю. М.* Анализ поэтического текста. Л., 1972.

100. *Лоэв М.* Теория вероятностей. М., 1962.

101. *Ляпунов А. А., Яблонский СВ.* О теоретических проблемах кибернетики.— В кн.: Кибернетика, мышление, жизнь. М., 1964.

102. *Мандельштам Л. И.* Лекции по оснонам квантовой механики. Полное собрание трудов, т. V. М., 1950.

103. *Марков А. А.* Распространение закона больших чисел на величины, зависящие друг от друга.— «Известия Казанского физико-математического общества», 1906, серия 2, т. 15,

104. *Марков А. А.* Избранные труды по теории чисел и теории вероятностей. Под редакцией 10. В. Вейника. М., 1951.

105. *Мельников Г. П.* Системный подход в лингвистике.— В кн.: Системные исследования. Ежегодник — 1972. М., 1972.

106. *Месарович М., Мако Д., Такахара И.* Теория иерархических многоуровневых систем. М., 1973.

107. *Мидлтпон Д.* Введение в статистическую теорию связи. М., 1961.

108. *Моран П.* Статистические процессы эволюционной теории. М., 1973.

109. *Налимов В. В.* Вероятностная модель языка. О соотношении естественных и искусственных языков. М., 1974.

110. *Налимов В. В.* Влияние идей кибернетики и математической статистики на. методологию научных исследований.— В кн.: Методологические проблемы кибернетики. (Материалы к всесоюзной конференции), т. .1. М., 1970.

111. *Налимов В. В., Мульченко Э. М.,* Наукометрия. М., 1969.

112. *Напалков А. В.* Кибернетика и пути изучения мозга.— В кн.: Кибернетика, мышление, жизнь. М., 1964.

113. *Напалков А. В.* Эвристическое программирование. Ростов-на-Дону, 1971.

114. *Напалков А. В., Чичварина П. А.* Можно ли моделировать работу мозга? М., 1966.

115. *Нейман Дж.* Вероятностная логика и синтез надежных организмов из ненадежных компонент.— В кн.: Автоматы. М., 1956.

116. *Нейман Дж.* Общая и логическая теория автоматов.— В кн.: *Тьюринг А.* Может ли машина мыслить? М., 1960.

117. *Николаев Л. А.* Основы физической химии биологических процессов. . М., 1971.

118. *Новик И. Б.* К характеристике взаимоотношений физики и кибернетики в свете диалектико-материалистического монизма.— В кн.: Методологические проблемы кибернетики. (Материалы Всесоюзной конференции). М., 1970.

119. *Новик И. Б.* Философские вопросы моделирования психики. М., 1969.

120. О положении в биологической науке. Стенограммы отчета сессии ВАСХНИЛ 31 июля — 7 августа 1948 г. М., 1948.

121. *Омельяновский М. Э.* Диалектика в современной физике. М., 1973.

122. *Опарин А. И.* Происхождение жизни на Земле. М., 1957.

123. *Орбели Л. А.* Об эволюционном принципе в физиологии.— В кн.: Вопросы эволюционной физиологии. М.— Л., 1961.

124. *Палиевский П. В.* Мера научности.— «Знамя», 1966, № 4.

125. *Паск Г.* Естественная история целей.— В кн.: Самоорганизующиеся системы. М., 1964.

126. *Паск Г.* Модель эволюции.— В кн.: Принципы самоорганизации. М., 1966.

127. *Патти Г.* Происхождение предбиологических систем. М., 1966.

128. *Петров С.* Субстрат, структура, свойства.— «Вопросы философии», 1958, № 10.

129. *Петрова Э.А.* Морфология серных пурпурных бактерий.— «Микробиология», 1959, т. 28. № 6.

130. *Петрушенко Л. А.* Самодвижение материи в свете кибернетики. Философский очерк взаимодействия организации и дезорганизации в природе. М., 1971.

131. *Пилипенко Н. В.* Взаимосвязь причинности, необходимости и случайности.— «Вопросы философии», 1973, № 9.

132. *Пирс Д.* Электроны, волны и сообщения. М., 1961.

133. *Позин И. В.* Моделирование нейронных структур. М., 1970.

134. *Пойа Д.* Математика и правдоподобные рассуждения. М., 1957.

135. *Пойа Д.* Как решать задачу. М., 1961.

136. *Полетаев И. А.* К определению понятия «информация».— В кн.: Исследования по кибернетике. М., 1970.

137. *Полушкин В. А.* К определению понятия «информация».— «Научно-техническая информация», 1963, № 9.

138. *Пригожин И.* Введение в термодинамику необратимых процессов. М., 1960.

139. *Пушкин В. Г.* Проблема надежности. Философский очерк. М., 1971.

140. *Равен X.* Оогенез. Накопление морфогенетической информации. М., 1966.

141. *Радушкевич Л. В.* Курс статистической физики. М., 1960.

142. *Разумовская 3. Т.* О виде микроорганизмов.— «Вестник ЛГУ, серия биологии», 1957, вып. 4, № 21.

143. *Рассел Б.* Человеческое познание. Его сфера и границы. М., 1957.

144. *Ревзин И. И.* К соотношению структурного и системного подходов в современной лингвистике.— В кн.: Системные исследования. Ежегодник — 1972, М., 1972.

145. *Рейтман У.* Познание и мышление. Моделирование на уровне информационных процессов. М., 1968.

146. *Ровенский З.И., Уемов А. И., Уемова Е. А.* Машина и мысль. М., 1960.

147. *Рогинский Л. Я., Левин М. Г.* Антропология. М., 1963.

148. *Розенблатт Т.Ф.* Принципы нейродинамики. Перцептроны и теории механизмов мозга. М., 1965.

149. *Рудь И. Д., Цукерман И. И.* Искусство и теория информации.— В кн.: Художественное и научное творчество. Л., 1972.

150. *Рунин В. М.* Творческий процесс в эволюционном аспекте.— В кн.: Художественное и научное творчество. Л., 1972.

151. *Самойлович А. Г.* Термодинамика и статистическая физика. М., 1955.

152. *Сачков Ю. В.* Введение в вероятностный мир. М., 1971.

153. *Сачков Ю. В.* Вероятность и развитие системно-структурных исследований.— В кн.: Системные исследования. Ежегодник — 1969. М., 1969.

154. *Сачков Ю. В.* Информация и вероятность.— «Вопросы философии», 1971, № 6.

155. *Свидерский Е. И.* О диалектике элементов и структуры в объективном мире и познании. М., 1962.

156. *Свидерский В. И., Зобов Р. А.* Уровни организации в свете представлений об элементах и структуре.— В кн.: Развитие концепции структурных уровней в биологии. М., 1972.

157. *Седов Е. А.* К вопросу о соотношении энтропии информационных процессов и физической энтропии.— «Вопросы философии», 1965, № 1.

158. *Сент-Дъердъи А.* Введение в субмолекулярную биологию. М., 1964.

159. *Серебрянников О. Ф.* Эвристические принципы и логические исчисления. М., 1970.

160. *Сетров М. И.* Организация биосистем. Л., 1971.

161. *Сетров М. И.* Степень и высота организации систем.— В кн. Системные исследования. Ежегодник — 1969. М., 1969.

162. *Смирнов В. А.* Основы радиосвязи на ультракоротких волнах. М., 1959.

163. *Смолуховский М.* О понятии случайности и о происхождении законов вероятностей в физике.— «Успехи физических наук», 1927, т. 7, вып. 5.

164. *Соколов Е. Н.* Механизмы памяти. М., 1969.

165. *Тарасенко Ф. П.* К определению понятия «информация» в кибернетике.— «Вопросы философии», 1963, № 4.

166. *Тахтаджян А. Л.* Тектология: история и проблемы.— В кн.: Системные исследования. Ежегодник — 1971. М., 1972.

167. *Терлецкий Я. П.* Статистическая физика. М., 1966.

168. *Тихомиров О. К.* Структура мыслительной деятельности человека. М., 1969.

109. *Том Р.* Динамическая теория морфогенеза.— В кн.: На пути к теоретической биологии. М., 1970.

170. *Тринчер К. С.* Биология и информация. Элементы биологической термодинамики. М., 1964.

171. *Тьюринг А. М.* Может ли машина мыслить? М., 1960.

172. *Тюхтин В. С.* Теория отражения в свете современной науки. М., 1971.

173. *Уилкс С.* Математическая статистика. М., 1967.

174. *Украинцев В. С.* Информация и отражение.— «Вопросы философии», 1963, № 2.

175. *Украинцев Б. С.* О возможностях кибернетики в свете свойства отражения материи.— В кн.: Философские вопросы кибернетики. М., 1961.

176. *Уоддингтон К. X.* Зависит ли эволюция от случайного поиска? В кн.: На пути к теоретической биологии. М., 1970.

177. *Уоддингтон К. X.* Морфогенез и генетика. М., 1964.

178. *Уоддингтон К. X.* Основные биологические концепции.— В кн.: На пути к теоретической биологии. М., 1970.

179. Управление, информация, интеллект. Под ред. А. И. Берга и др. М., 1976.

180. *Урсул А. Д.* Информация. Методологические аспекты. М., 1971.

181. *Ухтомский А. А.* Доминанта. Л., 1966.

182. *Файнстейн А.* Основы теории информации. М., 1960.

183. *Федоров Е. С.* Перфекционизм.— «Известия С.-Петербургской биологической лаборатории», т. 8. СПб., 1906.

184. *Фёрстер Г.* О самоорганизующихся системах и их окружении.—-В кн.: Самоорганизующиеся системы. М., 1964.

185. *Филдс У. С, Эббот У.* Предисловие.— В кн.: Концепция информации и биологические системы. М., 1066.

488. *Фогель Л., Оуэнс А., Уолш М.* Искусственный интеллект и эволюционное моделирование. М., 1969.

187. *Фок В. А.* Об интерпретации квантовой механики.— В кн. Философские проблемы современного естествознания. М., 1951)';

188. *Фролов Ю. П.* Диалектика живой природы и современная кибернетика.— В кн.: Философские вопросы кибернетики. М., 1961.

189. *Хайлов К. М.* Некоторые условия количественного подхода к организации биологических систем. В кн.: Системные исследования. Ежегодник — 1969. М., 1969.

190. *Хайлов К. М.* Упорядоченность биологических систем.— «Успехи биологических наук», 1966, т. 61, вып. 2.

191. *Харкевич А. А.* О ценности информации.— В кн.: Проблемы кибернетики, вып. 4. М., 1960.

192. *Харкевич А. А.* Очерки общей теории связи. М., 1955.

193. *Хартли Р.* Математическая теория связи.— В кн.: Теория информации и ее приложения. М., 1959.

194. *Хинчин А. Я.* Математические основания статистической механики. М.— Л., 1943.

195. *Хинчин А. Я.* Понятие энтропии в теории вероятностей.— «Успехи математических наук», 1952, т. VIII, вып. 3 (55).

196. *Хомский Н.* О некоторых формальных свойствах грамматик.— «Кибернетический сборник», вып. 5, 1962.

197. *Хомский Н.* Три модели описания языка.—«Кибернетический сборник», 1961, вып. 2.

198. *Хомский И. Миллер Д.* Введение в формальный анализ естественных языков.— «Кибернетический сборник», 1965, вып. 1.

199. *Цопф Г.* Отношение и контекст.— В кн.: Принципы самоорганизации. М., 1966.

200. *Чуковский К. И.* От двух до пяти. М., 1956.

201. *Шамбадаль П.* Развитие и приложение понятия энтропии М., 1967.

202. *Шеннон К. Е.* Бандвагон.— В кн.: Работы по теории информации и кибернетике. М., 1963.

203. *Шеннон К. Е.* Математическая теория связи.— В кн.: Работы по теории информации и кибернетике. М., 1963.

204. *Шеннон К. Е.* Надежные схемы из ненадежных реле.— В кн. *К. Шеннон.* Работы по теории информации и кибернетике. М. 1963.

205. *Шеннон К. Е.* Предсказание и энтропия английского печатного текста.— В кн.: К. *Шеннон.* Работы по теории информации и кибернетике. М., 1963.

206. *Шмальгаузен И. И.* Основы эволюционного процесса в свете кибернетики.— В кн.: Проблемы кибернетики, вып. 4. М., 1960.

207. *Штофф В. А.* Моделирование и философия. М.— Л., 1966.

208. *Шрейдер Ю. А.* О количественных характеристиках семантической информации.— «Научно-техническая информация», 1963, № 10.

209. *Шрейдер Ю.А.* Об одной модели семантической теории информации.— В кн.: Проблемы кибернетики, вып. 13. М., 1965.

210. *Эйген М.* Самоорганизация материи и эволюция биологических макромолекул. М., 1973.

211. *Эшби У. Росс.* Введение в кибернетику. М., 1959.

212. *Эшби У. Росс.* Конструкция мозга. Происхождение адаптивного поведения. М,, 1962,

213. *Эшби У. Росс.* Принципы самоорганизации.— В кн.: Принципы самоорганизации. М., 1966.

214. *Эшби У.* Росс. Схема усилителя мыслительных способностей,— В кн.: Автоматы. М., 1956.

215. *Юдин Г. А.* Диалектика части и целого. Алма-Ата, 1965.

216. *Юдин Э. Г.* Системная ориентация.— В кн.: Системные иссле-доппния. Ежегодник — 1972. М., 1973.

217. *Яглом А. М., Яглом И. М.* Вероятность и информация. М., 1973.

218. *Якобсон Р. О.* Типологические исследования и их вклад в сравнительно-историческое языкознание.— В кн.: Новое в лингвистике, вып. III. М., 1963.

219. *Ярошевский М. Г.* На путях к общей теории творчества.— В кн.: Художественное и научное творчество. Л., 1972.

220. *Aborn M., Rubinstein H.* Information theory and immediate recoil.— «Journal of Experimental Psychology», vol. 44, 1952.

221. *Ackoff R. L.* Towards a behavioral theory of communication.— «Management Science», 1958, vol. 4, N 3.

222. *Aubert V.* Chance in social affairs.—«Inquiry (University of Oslo)», 1959, N 2.

223. *Atteneave F.* Application of information theory to psychology: A summary of basic concepts, methods and results. N. Y., 1959.

224. *Atteneave F.* Symmetry, information and memory for patterns.— «American Journal of Psychology», vol. 68, 1955.

225. *Bain A.* The Senses and the Intellect. N. Y., 1874.

226. *Bancroft W. D.* A universal low.— «Science», vol. 33, 1911.

227. *Barber B.* Science and the social order. Glencoe, 1952.

228. *Bar-Hillel Y.* Semantic information and its measures.— In: Transactions of Tenth Conference on Cybernetics. N. Y., 1952.

229. *Bar-Hillel Y., Carnap R.* Semantic information.— «British Journal of Philosophical Science», 1953, vol. 4, N 14.

230. *Bar-Hillel Y., Carnap R.* An outline of theory of Semantic Information.— «Technical Report» (Massachusetts Institute of Technology), N 247, 1952.

231. *Bense M.* Einfiihrung in die informations theoretische Asthetik. Hamburg, 1969.

232. *Calvin M.* Chemical evolution and the origin of life.— «American Scientist», N 44, 1956.

233. *Cezan C.* La mode. Paris, 1967.

234. *Chomsky N., Halle M.* Sound pattern of English. N. Y., 1969.

235. *Cogniot G.* Le materialisme greco-romain. Paris, 1964.

236. *Diels H.* Die Fragmente der Vorsokratiker. Zurich, 1968.

237. *Elsasser W. M.* The physical foundation of biology. Ldn, 1958.

238. *Fishburn P. C.* On the prospects of a unified theory of value.— «Proceedings of Systems Science Conference». Cleveland, Ohio, 1965.

239. *Gamow G.* Possible relation between desoxyribonucleic acid and structure.— «Nature», N 173, 1954.

240. *Hick W. E.* On the rate of gain of information.— «The Quarterly Journal of Experimental Psichology», 1952, vol. 4, pt. 1.

241. *Howard P. A.* Information value . theory.—«Proceedings of Systems Science Conference». Cleveland, Ohio, 1965.

242. *Huxley I. S.* Soviet Genetics and World Science. London, 1949.

243. *Hyman R.* Stimulus information as a determinant of reaction times.— «Journal of Experimental Psychology», vol. 45, 1951!. N 3.

244. *Itany J.* Vocal Communication of the Wild Japanese Monkeys. «Primates», 1963, N 4.

245. *Jacobson R.* Selected Writings, vol. I. Gravenhage, 1962.

246. *Jaynes E. T.* Information theory and statistical mechanics. «Physical Review», vol. 106, 1957, N 5.

247. *Lamprecht S. P.* Contingency in nature.— «Philosophy anil Phenomenological Research», vol. 32, 1971, N 1.

248. *Lonard I. A.* Information Theory in Psychology. N. Y., 1956.

249. Mathematical challenges to the neo-Darwinan interpretation of evolution. Filadelfia, 1967 (Wistar Institute Symposium Monographs, N5).

250. *Miller G. A.* Information and memory — «Scientific American», vol. 195, 1956, N 2.

251. *Mills E.* The measurement of value of scientific information.— In.: Operation Research in Research and Development. New York — London, 1963.

252. *Newell A., Shaw J. G-, Simon II. A.* Elements of a theory of human problem solving.— «Psychology Review», N 65, 1958.

253. *Newell A., Shaw J. C, Simon H.A.* The process of creative thinking. University of Colorado, 1958 (Symposium on Cognation, Rand Corporation, Report P — 1320, 16 Sept. 1958).

254. *Nirenberg M. W., Matthaei I. H.* The dependence of cell-free protein synthesis in E — coli upon naturally occuring or synthetic polyrihonucleotides.— Proceedings of the National Academy of Science USA», vol. 47, N 10, 1961.

255. *Poincdre H.* Mathematical creation. The Foundations of Science. N. Y., 1913.

256. *Pringle I. W. S.* On the parallel between learning and evolution. «Behaviour», 1951, N 3.

257. *Skinner B. F.* Beyond freedom and dignity. N. Y., 1971.

258. *Schiller F. G. S.* Scientific discovery and logical proof.— In: Studies in the History and Method of Science. Oxford, 1917.

259. *Simpson G. G., Tiffany L. H., Pittendrig G. S.* Life. An introduction to biology. N. Y., 1957.

260. *Sneed Y. D.* Entropy, information and decision.— «Synthese», vol. 17, 1967, N 4.

Научно-практическое издание

Кононюк Анатолий Ефимович

Информациология

Общая теория информации

Книга 3

Авторская редакция

Подписано в печать 20.12.2010 г.
Формат 60х84/16.
Усл. печ. л. 16,5. Тираж 300 экз.

Издатель и изготовитель:
Издательство «Освита Украины»
04214, г. Киев, ул. Героев Днепра, 63, к. 40
Свидетельство о внесении в Государственный реестр
издателей ДК №1957 от 23.04.2009 г.
Тел./факс (044) 411-4397; 237-5992
E-mail: osvita2005@ukr.net, www.rambook.ru