МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ РСФСР

Ленинградский ордена Ленина
и ордена Трудового Красного Знамени
государственный университет

Ю.П.Петров

ИНФОРМАЦИЯ И ЭНТРОПИЯ В КИБЕРНЕТИКЕ

Учебное пособие

Ленинград 1989

Печатается по постановлению
Редакционно-издательского совета
Ленинградского университета

Петров Ю.П. Информация и энтропия в кибернетике: Учеб.пособие. Л., 1989.

В книге изложен генезис представлений об информации в системах управления и рассмотрена ее связь с энтропией физики. Рассмотренные положения применяются далее для количественных оценок информации в различных системах и прежде всего - в системах управления.

Рецензенты: д-р физ.-мат.наук, проф.В.Ф.Демьянов (Ленингр.ун-т),
канд.техн.наук, доц.В.Т.Шароватов (Ленингр.механ. ин-т)

© Ленинградский университет, 1989

Введение

В учебном пособии "Информация и энтропия в кибернетике" рассматриваются основные фундаментальные вопросы, лежащие в основе кибернетики.

Хорошо известно, что управление невозможно без информации. Но что такое информация? Достаточно ли нам житейского представления об информации как о некоторых знаниях, сообщениях? Согласуются ли с этими представлениями научные определения информации?

В научно-популярной литературе много раз встречается отождествление информации с отрицательной энтропией. Но что такое энтропия? Действительно ли она является мерой неупорядоченности? В чем сходство и в чем принципиальное различие между энтропией физики и энтропией распределения вероятностей в кибернетике? Можно ли отождествлять информацию с отрицательной энтропией? Вот тот круг вопросов, от ответа на которые зависит правильное понимание кибернетики и смежных с нею дисциплин – теории систем управления, передачи информации и т.д. Этот круг вопросов рассматривается в настоящей книге.

Книга начинается с анализа понятия энтропии и ее связи с обратимостью и необратимостью протекающих в природе процессов. Анализируется, какими путями возникло представление об энтропии как о мере неупорядоченности и показывается, почему это представление нуждается в существенном уточнении.

Рассмотрен также вопрос о справедливости принципа минимума производства энтропии в стационарном режиме открытой системы, поскольку этот принцип неоднократно использовался при анализе сложных систем управления. Использовать этот принцип надо осторожно, поскольку его некритическое применение может приводить к ошибкам.

Далее рассматривается понятие информации, анализируются различные определения информации, описываются методы вычисления количества информации. Помимо наиболее известной формы информации – статистической информации – дается представление о других интересных формах информации – целевой, облегчающей достижение какой-либо цели, а также смысловой информации, наиболее близко подходящей к тому, что обычно называют "знанием", "сообщением".

Подход к определению основных понятий кибернетики, развиваемый в книге, несколько отличается от того, который встречается на страницах научно-популярной литературы. Автор не хотел бы, чтобы это воспринималось как полемика, или упрек в адрес тех книг, где отождествлялась, например, информация с отрицательной энтропией. Такой подход на определенном этапе облегчал понимание и был, возможно, даже полезным. Однако сейчас настала пора для более строгого подхода. Автор надеется, что такой подход в конечном счете не усложнит, а облегчит понимание основных понятий кибернетики.

§ I. ЭНТРОПИЯ

Понятие энтропии вошло в физику в девятнадцатом веке, при анализе закономерностей превращения различных форм энергии в теплоту и теплоты - в работу. Само слово "энтропия" происходит от греческого слова "энтропо", что значит "превращение","изменение". Энтропия - одно из важнейших понятий физики, почти столь же важное и общее, как и понятие энергии. Действительно, одними из самых важных законов физики, с проявлениями которых приходится сталкиваться на каждом шагу - являются начала термодинамики, первое и второе начало.

Первое начало термодинамики - это закон сохранения энергии. Второе начало (если говорить коротко) - это закон возрастания энтропии.

Важность закона сохранения энергии общеизвестна. Он устанавливает, что в природе при любых взаимодействиях энергия не может ни исчезать, ни возникать из ничего, а может лишь переходить из одной формы в другую, причем во вполне определенных соотношениях.

Однако несмотря на всю важность закона сохранения энергии он не определяет направленности протекающих в природе процессов, не определяет путей превращения энергии. Направление природных процессов определяет второе начало термодинамики, в формулировке которого как раз и участвует понятие энтропии. Если говорить образно, то закон сохранения энергии - это "бухгалтер" природы, в то время как второе начало термодинамики - это ее "дирижер".

Представьте себе, что вы поставили кастрюльку с водой на горячую плиту. Закон сохранения энергии позволяет утверждать, что кастрюля получит ровно столько тепла, сколько потеряет плита, однако сам по себе закон сохранения не может еще предсказать, что тепло обязательно будет переходить от горячей плиты к холодной кастрюле - закону сохранения энергии не противоречит и обратный переход тепла от холодного тела к горячему, лишь бы холодная кастрюля потеряла тепла ровно столько, сколько получила горячая плита. Для того чтобы правильно предсказывать в каком направлении действительно будут протекать любые природные процессы, необходимо использовать второе начало термодинамики, а в формулировке второго начала как раз и участвует понятие энтропии.

Если абсолютная температура системы равна Т градусов и система получает извне некоторое малое количество тепла ΔQ (настолько малое, что температура системы не меняется), то приращение энтропии системы ΔS (в дальнейшем везде энтропию будем обозначать буквой S) будет равно

$$\Delta S = \frac{\Delta Q}{T}, \tag{I}$$

т.е. приращение энтропии равно приращению тепла, деленному на абсолютную температуру, при которой это приращение происходит.

Соотношение (I) является определением энтропии. Мы убеждаемся, что энтропия действительно является менее наглядным понятием, чем, например, понятие температуры. Энтропию нельзя ощутить или непосредственно измерить. Можно только измерить количество тепла, перешедшее от одной системы к другой при взаимодействии, измерить абсолютные температуры, а затем вычислять приращение энтропии по формуле (I). Причем еще нужно отметить, что формула (I) определяет энтропию только с точностью до постоянного слагаемого, что на практике, правда, не приводит к каким-либо трудностям, поскольку фактически всегда приходится иметь дело не с абсолютными значениями энтропии, а с разностью энтропий в различных состояниях системы.

Несмотря на малую наглядность понятия энтропии, что затрудняет первоначальное понимание, энтропия оказывается очень полезной при анализе направленности различных природных процессов. Так, с ее помощью легко объяснить, почему тепло всегда переходит от горячих тел к холодным, а не наоборот.

Напомним, что согласно второму началу термодинамики при любом процессе - в том числе и процессе теплопереноса - суммарная энтропия может только возрастать. Если количество тепла ΔQ переходит от тела с температурой T_1 к телу более холодному, с температурой T_2, где $T_2 < T_1$, то первое тело теряет энтропию $\Delta S_1 = \frac{\Delta Q}{T_1}$, а второе – получает $\Delta S_2 = \frac{\Delta Q}{T_2}$, результирующее изменение энтропии

$$\Delta S_2 - \Delta S_1 = \Delta Q \left(\frac{1}{T_2} - \frac{1}{T_1} \right) = \Delta Q \frac{T_1 - T_2}{T_1 T_2}$$

положительно и тем больше, чем больше разность температур T_1 и T_2. Если бы тепло переходило от холодного тела к более горячему, то результирующее изменение энтропии было бы отрицательным, что противоречит второму началу. Энтропия является, таким образом, характеристикой необратимости процессов теплопереноса.

Второе начало термодинамики позволяет также установить закономерности превращения тепла в работу. Пусть тепловая машина получает от нагревателя количество тепла Q_1 при температуре T_1; некоторая доля тепла превращается в работу, а остаток отдается холодильнику, температура которого равна T_2 (роль холодильника может играть и окружающая среда, атмосфера). Может ли тепло Q_2 равняться нулю, т.е. может ли тепло Q_1 нацело превращено в работу? Произведем подсчет изменения энтропии. Нагреватель, отдавая тепло Q_1, теряет энтропию $S_1 = \frac{Q_1}{T_1}$, холодильник приобретает $S_2 = \frac{Q_2}{T_2}$; так как результирующее изменение энтропии не может быть отрицательным, то

$$\frac{Q_1}{T_1} \leq \frac{Q_2}{T_2} \qquad (2)$$

и, следовательно,

$$\frac{Q_1 - Q_2}{Q_1} \leq 1 - \frac{T_2}{T_1} \quad . \qquad (3)$$

Неравенство (3) показывает, что отношение тепла, превращенного в работу $(Q_1 - Q_2)$ к общему количеству тепла Q_1, полученного от нагревателя, всегда меньше единицы. Некоторая часть тепла, равная Q_2 – не может быть превращена в работу и рассеивается в окружающем пространстве. Наивысший возможный коэффициент полезного действия преобразования тепла в работу может быть достигнут в идеализированной тепловой машине (работающей по так называемому циклу

Карно). В этом идеализированном случае неравенство (3) превращается в равенство, и мы получаем выражение для наибольшего возможного коэффициента полезного действия машины, работающей при постоянной температуре нагревателя T_1 и холодильника T_2 :

$$\eta_{макс} = 1 - \frac{T_2}{T_1} . \qquad (4)$$

Так, если паровая машина получает из котла пар с температурой 200°C (т.е. абсолютная температура пара равна 200+273=473°K), а абсолютная температура холодильника равна 300°K (т.е. 27°C), то коэффициенты полезного действия такой машины не может превысить

$$\eta_{макс} = 1 - \frac{300}{473} = 0,366$$

независимо от того, какие бы усовершенствования не вносились нами в машину.

Формула (4) подсказывает наиболее эффективный путь улучшения коэффициента полезного действия – повышение температуры рабочего тела, и техника давно идет по этому пути. Повышение температуры пара до 560°C (что сейчас уже достигнуто) позволяет отодвинуть теоретический предел коэффициента полезного действия до $\eta_{макс} = 1 - \frac{300}{833} = 0,64$.

Действительный коэффициент полезного действия, конечно, всегда ниже теоретически возможного.

Из неравенства (3) вытекает еще одно важное следствие. Умножив все члены неравенства (3) на Q_1 , получим

$$Q_1 - Q_2 \leq Q_1 - T_2 \frac{Q_1}{T_1} . \qquad (5)$$

В этом неравенстве член $Q_1 - Q_2$ соответствует тепловой энергии, которая может быть преобразована в другие виды – в механическую, электрическую и т.п. Эту энергию называют свободной энергией и обозначают буквой F . Член $\frac{Q_1}{T_1}$ в неравенстве (5) – это энтропия S , а Q_1 – полная энергия U ; мы убеждаемся, что только часть ее – F – свободна, а остальная часть – $T_2 \cdot S$ – это связанная энергия, осужденная к рассеянию в окружающей среде. Исключая из неравенства (5) знак равенства, возможный только для идеализированных систем, для всех реальных систем неравенство (5) можно записать в виде

$$F < U - T_2 S. \qquad (6)$$

Поскольку энтропия S может только возрастать, свободная энергия F может только убывать. Если мы будем рассматривать изолированную систему, не обменивающуюся энергией с другими системами, то согласно первому началу термодинамики, ее полная энергия U остается постоянной. В то же время согласно второму началу термодинамики ее свободная энергия F может только убывать. Все процессы в изолированной системе, любые преобразования энергии из одних форм в другие, могут идти только в направлении убывания свободной энергии F.

Таким образом, область применимости второго начала термодинамики не ограничивается тепловыми процессами. Оно утверждает, фактически, необратимость любых природных процессов. Любое взаимодействие, любое преобразование энергии их одних форм в другие, сопровождается переходом части энергии в тепло и необратимым увеличением энтропии. Энтропия выступает, таким образом, как характеристика необратимости всех природных процессов. Обратимые процессы являются лишь идеализацией. Таков вывод классической термодинамики.

Нужно, однако, иметь в виду, что методы классической термодинамики не используют представлений о молекулярном строении вещества, что неизбежно ограничивает глубину изучения, не позволяя вскрыть природу изучаемого явления.

Более глубокий подход возможен в рамках молекулярно-кинетической теории, когда тепловая энергия рассматривается как энергия беспорядочного движения отдельных молекул. Молекулярно-кинетическая теория применима к различным состояниям вещества, но наибольшее применение она получила при исследовании газов. Особенно просты выводы теории в применении к идеальному газу, когда молекулы газа можно рассматривать как идеально упругие бесконечно-малые шарики, взаимодействующие между собой только во время столкновений по законам упругого удара. Поведение реальных газов (особенно одноатомных благородных газов) с хорошей степенью точности подчиняется законам идеального газа.

Для идеального газа средний квадрат скорости молекул v^2_{cp}

(если в сосуде содержится n молекул и первая имеет скорость v_1, вторая - v_2 и т.д., то $v_{cp}^2 = \frac{1}{n} \sum_{i=1}^{n} v_i^2$) пропорционален абсолютной температуре. Температура оказывается, таким образом, мерой интенсивности движения молекул. Чем выше температура, тем быстрее (в среднем) движутся молекулы.

Молекулярно-кинетическая теория сталкивается с известными трудностями при объяснении необратимости природных процессов. Действительно, основное уравнение динамики, которому подчиняется движение каждой молекулы (как и любого тела) имеет вид

$$F = m \frac{d^2 S}{d t^2}, \tag{7}$$

где F - действующая на тело сила, m - его масса, $\frac{d^2 S}{d t^2}$ - ускорение. Время t входит в это уравнение в квадрате. Уравнения динамики не изменяются при замене t на $-t$, при изменении знака времени. Если сила F зависит только от положения тела (примерами таких сил могут служить силы тяготения, электростатического притяжения и отталкивания, а также силы, возникающие при упругом ударе), то симметричность уравнений динамики относительно знака времени приводит и к симметричности (относительно времени) протекания процессов, что должно приводить, казалось бы, к отсутствию всякой необратимости.

Действительно, если мы рассмотрим идеализированные, чисто механические системы, в которых отсутствует трение, а с ним и переход механической энергии в тепловую (маятник без трения в точке подвеса, идеально-упругий стальной шар, подпрыгивающий над массивной плитой), то в таких системах, состоящих из малого числа взаимодействующих тел, не удается заметить какой-либо необратимости.

Наиболее наглядно отсутствие необратимости проявляется, если процессы в системе заснять на кинопленку, а затем пустить фильм в обратную сторону. Для идеализированных механических систем без трения кинофильм, запущенный в обратную сторону не будет в чем-либо отличаться от нормального. Маятник в "кинофильме наоборот будет точно так же совершать колебания постоянной амплитуды, а стальной шар - подпрыгивать на одну и ту же высоту.

Необратимость процессов в реальных системах связана с наличием трения. При трении механическая энергия переходит в тепловую, появляется приращение тепла ΔQ, а с ним – и приращение энтропии $\Delta S = \frac{\Delta Q}{T}$, что свидетельствует о необратимости процесса. Однако если обратиться к непосредственному анализу движения молекул, то легко убедиться, что понятие трения к ним не применимо. При взаимодействии кинетическая энергия одной молекулы переходит в кинетическую энергию другой; трение – это понятие, относящееся ко взаимодействию макроскопических тел, а не отдельных молекул.

Рассмотрим идеальный газ, заключенный в левой половине сосуда с перегородкой (рис.I, в правой половине сосуда – пустота). Молекулы идеального газа взаимодействуют между собой и со стенками как идеально-упругие шары, трение исключается. Однако, если разрушить перегородку, опыт показывает, что газ заполняет обе половинки сосуда и не наблюдается возвращения к исходному состоянию, когда газ сосредоточен в левой половине сосуда. Налицо необратимый процесс.

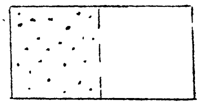

Рис.I

Как связать необратимость молекулярно-кинетических процессов с симметричностью относительно времени уравнений движения, которым они подчиняются? Движение каждой молекулы в отдельности, безусловно, полностью обратимо. Каким же образом из совокупности обратимых каждое в отдельности движений складывается необратимый в целом процесс расширения газа?

Для ответа на этот вопрос нам придется обратиться к представлениям теории вероятностей.

Рассмотрим идеальный газ, заключенный в сосуд. Молекулы газа двигаются как идеально-упругие шары с большими линейными скоростями, беспрестанно сталкиваясь как друг с другом, так и со стенками сосуда. В этих условиях траектории движения молекул имеют вид чрезвычайно сложных ломаных линий.

Разделим сосуд мысленно на равные ячейки. В беспорядочном движении молекул газа в каждой ячейке будут в разные моменты времени оказываться разные числа молекул. Однако не все эти числа будут равновероятными. Очевидно, что наблюдая за движением молекул, мы большую часть времени будем видеть более вероятные числа молекул и меньшую часть – менее вероятные. Сравним теперь между собой вероятности различных распределений молекул по ячейкам. Очевидно, что вероятность того, что некоторая произвольно выбранная молекула окажется в ячейке с номером i равна $\frac{1}{z}$, где z – число ячеек. Вероятность того, что в первой ячейке окажется ровно n_1 молекул, во второй – n_2, в третьей – n_3 и так далее будет равна, согласно известной формуле теории вероятностей

$$P = \frac{n!}{n_1!\, n_2! \ldots n_z!} \cdot \frac{1}{z^n}, \tag{8}$$

где n – общее число молекул в сосуде. Поясним эту формулу на примерах. Пусть в сосуде находится 10 молекул и он разделен на две равные ячейки. Тогда вероятность того, что все молекулы соберутся в правой ячейке, а левая останется пустой, равна

$$P_{0,10} = \frac{10!}{0!\,10!} \cdot \frac{1}{2^{10}} = \frac{1}{1024} \tag{9}$$

(поскольку $0! = 1$).

Аналогично вероятность того, что в правой ячейке будет 9 молекул, а в левой – одна, равна

$$P_{1;9} = \frac{10!}{1!\,9!}\,\frac{1}{2^{10}} = \frac{10}{1024}$$

Точно так же можно вычислить, что

$$P_{2,8} = \frac{45}{1024}; \quad P_{3;7} = \frac{120}{1024}; \quad P_{4,6} = \frac{210}{1024};$$

$$P_{5,5} = \frac{252}{1024}, \quad P_{6,4} = \frac{210}{1024}; \quad P_{7;3} = \frac{120}{1024};$$

$$P_{8,2} = \frac{45}{1024}, \quad P_{9,1} = \frac{10}{1024}, \quad P_{10;0} = \frac{1}{1024}.$$

Это различие в вероятностях связано с тем, что, например, состояние, когда все молекулы сосредоточены в правой половине сосуда, может быть осуществлено одним единственным расположением молекул (микросостоянием), а именно когда все молекулы – от первой до десятой – расположены справа. В то же время, например, состояние, в котором справа находится 9 молекул, а слева – одна, может быть реализовано десятью возможными расположениями (микросостояниями). Первым микросостоянием можно, например, считать такое, когда молекула с условным номером один находится в правой половине, а остальные – в левой, вторым состоянием – когда справа находится молекула с условным номером два, а остальные – в левой половине – и т.п. Точно также состояние, когда и в левой и в правой частях сосуда находится по 5 молекул может быть осуществлено 252 различными микросостояниями. Если число микросостояний обозначить W, то $W_{0,10} = 1$; $W_{1,9} = 10$; $W_{5,5} = 252$ и т.п. Число микросостояний, соответствующих данному состоянию, называют иногда "термодинамической вероятностью". Это не совсем точно, ибо с математической точки зрения вероятность есть отношение числа микросостояний, соответствующих данному состоянию системы к общему числу возможных микросостояний (в нашем примере – 1024). Это надо иметь в виду, что в математике вероятность всегда является числом, не большим единицы.

Наибольшее число микросостояний, а с ним и наибольшая вероятность соответствует состоянию, когда молекулы идеального газа равномерно распределены по ячейкам. В нашем примере это состояние в 252 раза вероятнее состояния, когда все молекулы сосредоточены в правой ячейке. Эта разница очень резко возрастает с увеличением общего числа молекул в сосуде.

Так, если в сосуде всего 4 молекулы, то вероятность их равномерного распределения (две – в правой половине сосуда и две – в левой), всего в 6 раз больше вероятности того, что все молекулы сосредотачиваются в правой ячейке; действительно,

$$P_{2,2} = \frac{4!}{2!\,2!} \cdot \frac{1}{2^4} = \frac{6}{16}\,; \quad P_{0;4} = \frac{4!}{0!\cdot 4!} \cdot \frac{1}{2^4} = \frac{1}{16}\,.$$

При 10 молекулах в сосуде вероятности различаются уже в 252 раза, при 16 молекулах – более чем в 12 тыс.раз, при 20 молекулах – более чем в 184 тыс.раз.

Для больших чисел молекул n_i формулу для вычисления вероятности распределения молекул по ячейкам (8) можно записать в другом виде, учитывая, что для больших n с хорошей степенью точности выполняется приближенное равенство:

$$\ln n! = n(\ln n - 1)\,. \tag{10}$$

Используя это равенство, логарифм вероятности распределения молекул по ячейкам можно записать в виде:

$$\ln P = n(\ln n - 1) - \sum_{i=1}^{z} n_i \ln n_i + \sum_{i=1}^{z} n_i - n \ln z. \qquad (11)$$

Поскольку общее число молекул n и число ячеек заданы, и $\sum_{i=1}^{n} n_i = n$ то можно записать

$$\ln P = C_1 - \sum_{i=1}^{z} n_i \ln n_i, \qquad (12)$$

где C_1 - постоянная.

Аналогично можно записать:

$$\ln W = C_2 - \sum_{i=1}^{z} n_i \ln n_i, \qquad (13)$$

где C_2 - другая постоянная.

Сумма $\sum_{i=1}^{z} n_i \ln n_i$, как это легко установить средствами дифференциального исчисления, минимальна тогда, когда все n_i равны друг другу. Для примера сравним два распределения тысячи молекул по двум равным ячейкам: первое, где 400 молекул находятся в одной ячейке и 600 - в другой, и второе, когда в каждой ячейке находятся по 500 молекул; для первого распределения

$$\sum_{i=1}^{2} n_i \ln n_i = 400 \ln 400 + 600 \ln 600 = 6236{,}7,$$

а для второго

$$\sum_{i=1}^{2} n_i \ln n_i = 500 \ln 500 + 500 \ln 500 = 6214{,}6.$$

Разность логарифмов обоих распределений равна 22,1, а это означает, что второе распределение в $e^{22,1} \approx$ 3,6 миллиарда раз вероятнее первого.

При тех колоссальных количествах молекул с которыми приходится иметь дело в реальных объемах газа (так, в одном литре кислорода при атмосферном давлении содержится 2,7 10^{22} молекул) различие между вероятностями различных состояний чрезвычайно велико и практически будут наблюдаться только вероятные состояния, с равномерным распределением молекул газа по объему.

Если маловероятное состояние создано искусственно (например, перегородкой, отделяющей левую половину сосуда – с повышенным давлением газа – от правой), то после удаления перегородки будет происходить переход к более вероятному состоянию с равномерным распределением молекул по объему.

В то же время известно, что при расширении идеального газа возрастает его энтропия. Таким образом, вероятность и энтропия изменяются в одном направлении, и можно считать, что энтропия есть функция вероятности

$$S = f(W). \tag{I4}$$

Если рассмотреть систему, состоящую из двух частей, то энтропия ее равна сумме энтропий составляющих частей (энтропии складываются), а вероятность состояния системы равна произведению вероятностей состояния составляющих систему частей. Исходя из этого (мы не будем останавливаться на деталях доказательства) формулу (I4) можно записать в виде

$$S = k \ln W, \tag{I5}$$

т.е. энтропия пропорциональна логарифму вероятности состояния. Формула (I5) – это знаменитая формула Л.Больцмана. Постоянная в ней – это универсальная постоянная (постоянная Больцмана), связывающая энтропию и вероятность. Для численного вычисления постоянной Больцмана, достаточно вычислить S и W один раз для простейшего случая (для идеального газа); это вычисление было произведено Больцманом, который установил, что в формуле (I5)

$$k = 1{,}38\ 10^{-16}\ \frac{эрг}{градус} = 3{,}296 \cdot 10^{-24}\ \frac{кал}{градус}. \tag{I6}$$

Вообще говоря, надо еще отметить, что в формуле (I5) фигурирует не математическая вероятность p, всегда меньшая единицы, а термодинамическая вероятность W, отличающаяся от p на постоянный множитель. Однако, так как на практике всегда приходится иметь дело не с абсолютным значением энтропии S, а с разностью энтропий двух состояний S_1 и S_2, и, следовательно, с логарифмом частного $\ln \frac{W_1}{W_2}$, то поскольку $\ln \frac{W_1}{W_2} = \ln \frac{p_1}{p_2}$ (I7) различие между термодинамической и математической вероятностью не играет роли.

Процесс выравнивания концентраций газа в двух половинах сосуда после удаления перегородки (процесс, сопровождающийся увеличением энтропии) получает свое объяснение как переход от маловероятного состояния к более вероятному. Обратный процесс концентрации всех молекул в одной половине сосуда хотя и не противоречит законам столкновения между молекулами, однако имеет исключительно малую вероятность и поэтому на практике никогда не наблюдается.

Тем же переходом от маловероятных состояний к более вероятным объясняется и выравнивание температуры у обменивающихся теплом тел, переход тепла от горячих тел к холодным. Так, если приведены в соприкосновение горячее и холодное тело, то мы имеем маловероятное состояние, когда большая часть молекул с большими скоростями сосредоточены в одной половине системы, а в другой собраны в основном более медленные молекулы. При наличии переноса тепла между телами произойдет переход к гораздо более вероятному состоянию, когда пропорция быстрых и медленных молекул в обоих телах одинакова, что эквивалентно равенству их температуры.

Формула (I5) позволяет оценить вероятности различных отклонений температуры от равновесия. Пусть приведены в соприкосновение два объема воды по I см^3 каждый, температурой 27°С (300°К). Вычислим вероятность самопроизвольного перехода системы из состояния равновесия в такое состояние, когда температуры соприкасающихся объемов воды различаются на одну сотую градуса. Такой переход вызывает уменьшение энтропии на

$$\Delta S = \frac{0{,}01}{300^\circ} = \frac{1}{3}\ 10^{-4}\ \frac{\text{кал}}{\text{град}}$$

В то же время согласно формуле (I5),

$$\Delta S = k \ln \frac{P_1}{P_2}$$

где P_1 – вероятность исходного, равновесного состояния, а P_2 – вероятность нового состояния. Следовательно,

$$\frac{P_1}{P_2} = e^{\frac{\Delta S}{k}} = e^{\frac{0,33}{3,296}\cdot\frac{10^{-4}}{10^{-24}}} \approx e^{19} \approx 10^{16},$$

т.е. вероятность перехода хотя бы 0,0I кал от холодного тела к горячему невообразимо мала и на практике никогда не будет наблюдаться. Однако переходы очень малых количеств тепла (тепловые флюктуации) вполне возможны. Так если ΔQ = I0^{-2I} кал, то $\frac{P_1}{P_2} = e^{\frac{0,33}{3,296}\cdot\frac{10^{-23}}{10^{-24}}} \approx 2,7$ т.е. такой переход вполне вероятен (и мог бы наблюдаться с помощью прибора, улавливающего изменения температуры до I0^{-2I} градуса).

Поведение этой системы можно наглядно изобразить на графике, где по оси ординат отложена разность температур двух объемов, а по оси абсцисс - время (рис.2). Нужно только иметь в виду, что масштабы рисунка носят условный характер.

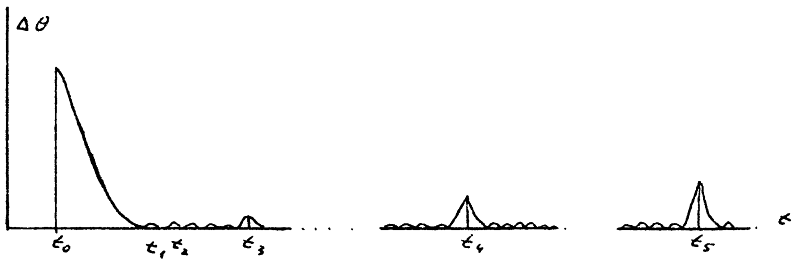

Рис.2

Пусть до момента времени t_0 объемы с разной температурой разделены непроницаемой перегородкой. Сразу после удаления перегородки идет (сопровождающийся увеличением энтропии) процесс выравнивания температур, температурная разность быстро уменьшается. Однако, она не остается равной нулю. Вокруг равновесного состояния существуют все время небольшие флюктуации (порядка I0^{-20}÷I0^{-2I} градуса). На рисунке такие флюктуации показаны в моменты времени t_1 ; t_2 . Изредка будут наблюдаться чуть более значительные флюктуации (момент времени t_3). Если ждать очень долгое время (миллионы лет), то можно дождаться несколько более значительных флюктуаций (моменты времени t_4 , t_5 , показанные, конечно, не в масштабе), но в общем система подавляющую часть времени будет находиться в состояниях очень близких к состоянию равновесия (максимума энтропии) и практически неотличимых от него.

Таким образом, кинетическая теория материи позволяет уточнить выводы классической термодинамики, позволяет понять, что второй закон термодинамики является вероятностным законом. Он применим лишь к системам, состоящим из очень большого числа взаимодействующих частиц. При тех колоссальных числах молекул, с которыми приходится иметь дело в реальных макроскопических телах отклонения от закона возрастания энтропии практически не наблюдаются, однако нужно всегда иметь в виду существование флюктуаций вокруг состояния равновесия. Малые флюктуации происходят все время, чем больше флюктуация, тем меньше вероятность ее появления, и при большом числе молекул сколько-нибудь значительная флюктуация практически никогда не будет появляться хотя вероятность ее появления не равна в точности нулю, а только очень, очень мала.

Кинетическая теория позволяет уточнить понимание энтропии. Энтропия системы пропорциональна логарифму вероятности состояния системы, т.е. является мерой вероятности состояния. Изолированная система, находящаяся в маловероятном состоянии (т.е. вдали от максимума энтропии) и предоставленная самой себе, перейдет со временем в более вероятное состояние и энтропия ее возрастет. В этом и заключается второе начало термодинамики для изолированных систем.

Рассмотрим теперь те возражения, которые выдвигались в свое время против кинетической теории Больцмана. Одно из таких возражений было сделано Лошмидтом и касалось возможности обращения скоростей молекул.

Если в некоторый момент времени t_0 система находилась в состоянии, далеком от термодинамического равновесия (например, в левой половине сосуда - газ, в правой - пустота (см.рис.I), или приведены в соприкосновение два тела - одно холодное, другое горячее), то по истечении некоторого времени через ряд промежуточных состояний установится состояние термодинамического равновес я. Если в один из моментов термодинамического равновесия изменить направление вектора скорости каждой молекулы на обратное (что на практике сделать, конечно, невозможно, но теоретически вполне мыслимо), то каждая молекула в отдельности, а следовательно, и система в целом пройдут все промежуточные состояния в обратном направлении и через время $t_1 - t_0$ система снова окажется в исходном, мало вероятном термодинамическом состоянии – состоянии с малой энтропией, в котором она находилась в момент времени t_0 (рис.2). Из этого обстоятельства Лошмидт делал вывод о том, что эволюция системы может проходить как в направлении возрастания, так и в направлении убывания энтропии и законы механики не могут объяснить необратимости естественных процессов.

Это рассуждение (известное в истории как "парадокс Лошмидта") вызвало многочисленные дискуссии и способствовало более точной формулировке основ кинетической теории материала. Простейшее возражение состоит в том, что идеально-точное обращение вектора скорости всех молекул невозможно, а уже самое малое отклонение приведет к тому, что система не пройдет в точности обратную последовательность состояний и не вернется в исходное состояние. Это возражение отрицает фактически правомерность "мысленного" (идеализированного) эксперимента. Однако мысленные эксперименты и идеализированные системы играют большую роль при построении многих физических теорий и имеют полное право на существование. Если мы не можем объяснить мысленного эксперимента, не противоречащего законам физики, то это свидетельствует лишь о неполном понимании сущности рассматриваемого явления. Парадокс Лошмидта получает разрешение, если мы проследим за дальнейшей судьбой системы, вернувшейся (после обращения вектора скорости всех молекул) в исходное состояние. Рассмотрим систему, поведение которой изображено на рис.2 и в момент времени t_1 изменим на обратные скорости всех молекул. После такого обращения система начнет проходить свои состояния в обратном порядке (рис.3) и через время $t_1 - t_0$ окажется в начальном состоянии t_0 . Однако это не значит, что она в нем останется. Поскольку разделяющей нетеплопроводной перегородки нет, система из состояния t_0 быстро перейдет в более вероятные состояния с ничтожно-малыми разностями температур и будет в дальнейшем находиться в них, испытывая лишь малые флюктуации вокруг состояния равновесия. Обращение скоростей всех молекул привело, как мы видим, лишь к тому, что неравновесное состояние (соответствующее моменту t_0) промелькнуло, как кратковременная флюктуация, на фоне вероятных состояний, в которых пребывает система подавляющую часть времени. Но такой же флюктуации можно было дождаться и без обращения скоростей молекул, если ждать колоссально-долгое время (многие миллиарды миллиардов лет). Идеализированный опыт обращения скоростей молекул приводит лишь к теоретически возможному броску по оси времени в окрестность маловероятного состояния, но не может отменить того факта, что система подавляющую часть времени будет находиться в вероятных состояниях, в окрестностях состояния равновесия, а большие самопроизвольные флюктуации, сопровождающиеся существенным уменьшением энтропии, являются событиями настолько редкими, что хотя они теоретически возможны (и на интервалах в миллиард миллиардов и более лет непременно осуществятся), но на практике при тех же интервалах времени, с которыми фактически приходится иметь дело никогда не наблюдались и не наблюдаются.

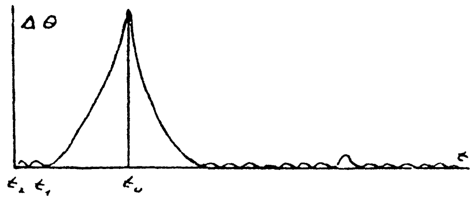

Рис.3

Другим интересным возражением против кинетической теории Л.Больцмана было возражение Цермело. Это возражение опирается на теорему Пуанкаре, справедливую для системы с любым числом частиц. Эта теорема утверждает, что какое бы состояние системы мы не взяли (хотя бы и маловероятное), по истечении достаточно длинного времени система обязательно повторит это состояние (либо, во всяком случае, будет отличаться от него на бесконечно малую величину ε). Так, если мы рассмотрим снова систему, поведение которой изображено на рис.2, то теорема Пуанкаре позволяет утверждать, что с течением времени система будет неограниченное число раз находиться и в состоянии, имевшем место в момент времени t_0 , так же, как и в состояниях, имевших место в моменты времени t_1; t_2;. ,

t_5 и т.д. Необходимо лишь наблюдать за системой достаточно долгое время.

Из теоремы Пуанкаре вытекает, что на бесконечно длинном интервале времени будут одинаково часто наблюдаться как переходы от менее вероятных состояний к более вероятным, так и обратные переходы – от более вероятных к менее вероятным состояниям, то есть от состояний с большой энтропией к состояниям с малой энтропией. Получается, что на достаточно длинном интервале времени должны одинаково часто наблюдаться как процессы увеличения энтропии, так и процессы уменьшения энтропии. На первый взгляд, это полностью противоречит второму началу термодинамики.

Теорема Пуанкаре доказана математически строго и сомневаться в ее справедливости нет оснований. Парадокс Пуанкаре–Цермело разрешается, если мы оценим время, необходимое для того, чтобы система, находящаяся в одном из маловероятных состояний, вернулась в это состояние. Пусть, например, в воздухе при атмосферном давлении выделена шаровая поверхность с радиусом I см. Если, в начальный момент времени концентрация кислорода внутри поверхности отличалась от окружающей на I%, то после неизбежного выравнивания концентраций газ, согласно теореме Пуанкаре, снова рано или поздно, вернется в начальное состояние, но это произойдет (согласно вычислениям М.Смолуховского), в среднем более чем через $10^{10^{13}}$ лет! Такова длительность "цикла Пуанкаре" для данного случая. Для промежутков времени более коротких чем "циклы Пуанкаре", мы не можем ожидать самопроизвольного появления флюктуации. Если она наблюдалась в начальный момент времени как следствие действия внешних по отношению к данной системе сил (например – вследствие наличия перегородки, разделяющей правую и левую половинки сосуда), то после удаления перегородки на любых реально возможных отрезках времени мы будем наблюдать необратимый процесс выравнивания концентраций газа, хотя теоретически, если подождать $10^{10^{13}}$ и более лет, то можно дождаться и момента, когда весь газ соберется в одной половине сосуда.

Парадокс Пуанкаре–Цермело не опровергает второго начала термодинамики, но уточняет его, выявляет то, что само понятие обратимого и необратимого процесса является понятием относительным. Процесс, необратимый в некоторой системе на одном участке времени, может оказаться обратимым в другой системе, или на другом (более длинном) временном отрезке. Так, например, М.Смолуховским был произведен расчет времени обратимости для различных объемов воздуха. "Представим себе, - пишет М.Смолуховский, - что в атмосферном воздухе нормальной плотности проведена шаровая поверхность радиуса r и поставим перед собой следующий вопрос: по истечении какого времени можно ожидать такого самопроизвольного нарушения нормального состава воздуха, чтобы концентрация кислорода в шаре оказалась на 1% выше нормальной " Приведем конечный результат расчета: при радиусе $r = 10^{-5}$ см время возврата $t = 10^{-11}$ с; при $r = 2{,}5 \cdot 10^{-5}$ см, $t = 1$ с; при $r = 3 \cdot 10^{-5}$ см, $t = 10^6$ с $\approx$ 12 суток и, наконец, при $r = 1$ см, $t = 10^{(10^{14})}$ с или больше чем $10^{(10^{13})}$ лет.

Таким образом, если наблюдать за объемом газа в шаре радиусом $3 \cdot 10^{-5}$ см, то при длительности наблюдения 5÷10 минут процесс будет казаться необратимым, а на интервале времени в десятки суток он будет обратим. Наконец, при $r = 1$ см практически уже никогда, при любых реально мыслимых сроках наблюдения не будет наблюдаться обратимость процесса, хотя и здесь предельный переход к $t \to \infty$ недопустим.

Понимание ограниченности второго начала термодинамики, понимание того, что обратимость или необратимость процесса тесно связаны с длиной интервала времени, на котором протекает процесс - весьма важно. Действительно, еще в середине XIX века на заре развития термодинамических методов Рудольф Клаузиус, абсолютизируя второе начало термодинамики, пришел к выводу о неизбежности "тепловой смерти" вселенной. Распространяя второе начало на всю бесконечную вселенную, Клаузиус считал, что рано или поздно энтропия вселенной достигнет максимума, свободная энергия упадет до нуля, температуры всех тел выравнятся, после чего исчезнет всякая возможность дальнейших изменений и превращений, и вселенная перейдет в состояние неподвижности и застоя, которые Клаузиус образно назвал состоянием "тепловой смерти". Этот вывод Клаузиуса получил (если можно так выразиться) "незаслуженную популярность"; он неоднократно повторялся и в XX веке. Так, Джинс писал в 1930 г.: "Вселенная не может существовать вечно, рано или поздно должно наступить время, когда последний эрг энергии достигнет наинизшей ступени на лестнице падающей полезности и в этот момент активная жизнь вселенной должна будет прекратиться". Церковь использовала представления о "тепловой смерти" вселенной как доказательство необходимости существования бога. Действительно, если следуя Р.Клаузиусу признать, что в природе возможны только процессы, ведущие к термодинамическому равновесию, а в настоящее время мир не находится в равновесном состоянии, то это значит, что в какой-то момент в прошлом мир был либо создан, либо получил от внешней силы начальный толчок, последствия которого мы и наблюдаем сегодня.

Представления Р.Клаузиуса о неизбежности "тепловой смерти" вселенной были подвергнуты резкой критике Ф.Энгельсом в "Диалектике природы". Энгельс критиковал взгляды Р.Клаузиуса с философских позиций, указывая на недопустимость безоговорочной экстраполяции законов, справедливых для ограниченного круга явлений, на всю вселенную. Теорема Пуанкаре конкретизирует для частного случая обратимости процессов общее философское положение об ограниченности возможности безусловной эстраполяции; процесс, необратимый на одном отрезке времени, может оказаться полностью обратимым при расширении интервала времени наблюдения. Второе начало термодинамики не относится к тем законам природы, которые можно экстраполировать на сколь угодно большие интервалы пространства и времени и теория "тепловой смерти" вселенной является необоснованной.

§ 2. ЯВЛЯЕТСЯ ЛИ ЭНТРОПИЯ МЕРОЙ НЕУПОРЯДОЧЕННОСТИ?

Рассмотрев подробно характеристики и свойства энтропии, мы имеем теперь возможность судить - является ли справедливой широко распространенная трактовка энтропии как меры неупорядоченности (или меры беспорядка, меры хаоса и т.п.), и понимание второго начала термодинамики как закона о неминуемом возрастании беспорядка в изолированной системе.

Трактовка энтропии как меры беспорядка весьма распространена. Ее можно встретить в монографиях, некоторых учебниках, в популярных книгах. Так, например, в монографии М.В.Волькенштейна "Молекулы и жизнь" (изд. "Наука", 1965) на с.30 читаем: "согласно второму началу термодинамики упорядоченность изолированной системы должна убывать, при соответствующем возрастании меры неупорядоченности – энтропии".

В популярной книге Б.В.Ахлибинского и Н.И.Храленко "Чудо нашего времени (кибернетика и проблемы развития)", Лениздат, 1963, на с.48 встречаем аналогичное утверждение: "энтропию можно рассматривать как меру беспорядка в движении молекул. Второе начало термодинамики, которое утверждает, что в изолированной системе энтропия возрастает, получает тот смысл, что в изолированной системе увеличивается беспорядок в движении молекул".

Число подобных утверждений можно легко умножить, поскольку убеждение в том, что энтропия является мерой беспорядка (или неупорядоченности, хаотичности и т.д.) действительно весьма широко распространено. Однако попробуем проанализировать более детально – в какой мере обосновано подобное убеждение? Действительно ли второе начало термодинамики утверждает неминуемость возрастания беспорядка, хаотичности во всех реально протекающих в природе процессах, поскольку все реальные процессы кроме идеализированных случаев, являются необратимыми и сопровождаются увеличением энтропии?

Для того чтобы придать доказательность и определенность дальнейшим рассуждениям, нам надо выбрать достаточно четкое определение понятия "беспорядок". В выборе определения мы будем следовать профессору А.И.Китайгородскому, который в своей известной книге "Порядок и беспорядок в мире атомов" (Физматгиз, 1959) определяет беспорядок через изотропность, отсутствие выделенности, предпочтительности отдельных областей пространства, отдельных направлений и т.д. Так, если речь идет о беспорядке в пространственном расположении частиц, то наиболее беспорядочным будет расположение, когда отсутствуют участки повышенной концентрации частиц и они примерно равномерно распределены по всей области, подчиняясь лишь законам случая. Например, на рис.4 изображено более упорядоченное расположение частиц (все они сосредоточены в левой половине сосуда), а на рис.5 – менее упорядоченное, частицы распределены по всему сосуду. Аналогично если речь идет о беспорядке в направлениях движения, то более упорядоченным будет состояние, когда преобладают движения одного направления, и менее упорядоченным - состояние, когда равновероятны любые значения вектора скорости. Рис.6 и 7 изображают соответственно упорядоченное и неупорядоченное состояние (стрелки условно показывают направление вектора скорости). Определение А.И.Китайгородского соответствует нашим интуитивным представлениям о порядке и беспорядке.

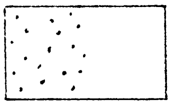

Рис.4

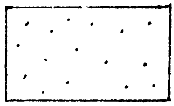

Рис.5

Если теперь, опираясь на приведенное определение неупорядоченности, рассмотреть происходящие в природе естественные необратимые процессы, то действительно можно найти большое количество примеров, когда увеличение энтропии в необратимом процессе сопровождается увеличением беспорядка. Так, если рассматривать рис.4 как условное изображение молекул газа, заключенных в левой половине сосуда, то после удаления удерживающей перегородки произойдет процесс расширения газа и он перейдет в состояние, условно изображенное на рис.5. Аналогично столкновения частиц между собой и со стенками сосуда быстро переведут упорядоченную систему, изображенную на рис.6, в неупорядоченное состояние - рис.7. И здесь увеличение энтропии в необратимом процессе взаимодействия между частицами приводит и к увеличению неупорядоченности. Большое число подобного рода примеров, возможно, и послужило основанием для утверждения о том, что энтропия есть мера неупорядоченности системы. Однако для того чтобы энтропия действительно могла служить такой мерой, недостаточно любого числа подкрепляющих примеров. Необходимо, чтобы не существовало ни одного опровергающего примера, ни одного случая, когда возрастание энтропии в каком-либо необратимом процессе в изолированной системе сопровождалось бы не уменьшением, а ростом упорядоченности, ни одного случая, когда установившееся состояние изолированной сис-

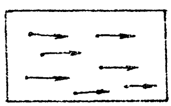

Рис.6

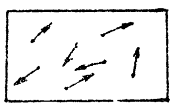

Рис.7

темы, соответствующее максимуму энтропии, в то же время не было бы состоянием с наибольшей неупорядоченностью.

Существование же таких случаев обнаруживается, когда мы начинаем рассматривать системы, в которых помимо сил отталкивания между молекулами действуют иные силы - силы тяготения, электрические и магнитные силы, силы химического взаимодействия и т.п.

Рассмотрим газ, находящийся в поле тяготения Земли (земную атмосферу). Равновесное распределение газа по высоте, распределение, соответствующее максимальной энтропии, отнюдь не будет наиболее беспорядочным распределением. Плотность газа будет убывать с высотой (примерно по экспоненциальному закону), большая часть газа (атмосферы) будет сосредоточена на сравнительно малом расстоянии от Земли. Если какой-либо внешней силой добиться наиболее беспорядочного распределения молекул по высоте (например, добиться равномерной плотности на высотах от 0 до 10 км), то после удаления внешней силы произойдет необратимый процесс восстановления первоначального, экспоненциального расщепления по высоте. В этом процессе энтропия возрастает, беспорядок уменьшается.

Могут возразить, что мы рассматриваем не изолированную систему, рассматриваем газ, находящийся во внешнем поле силы тяжести. Но это не так. В действительности мы рассматриваем изолированную систему - Земля (и ее поле тяготения) плюс земная атмосфера. В этой изолированной системе устанавливается распределение молекул, соответствующее максимуму энтропии, и оно не будет совпадать с максимальным беспорядком.

Можно рассмотреть также случай, когда поле тяготения порождается самими молекулами. Достаточно представить себе скопление межзвездного газа, достаточно большое для того, чтобы могли проявиться силы тяготения, и мы убедимся, что эти силы превращают скопление в довольно упорядоченную туманность, где плотность убывает от центра к периферии. В астрофизике строго доказывается, что наиболее неупорядоченное распределение газа, характерное примерно постоянной плотностью для всего пространства, не устойчиво. Силы тяготения приводят к тому, что газ самопроизвольно и необратимо концентрируется в упорядоченные скопления - зародыши будущих звезд. Самопроизвольный переход от упорядоченного состояния - рис.4 - к неупорядоченному – рис.5 – происходит лишь в малых объемах газа, где силы тяготения пренебрежимо малы.

Электрические и магнитные силы также могут упорядочивать частицы. Если мы представим себе множество стрелок (рис.8), способных свободно вращаться на оси, на которые действуют только случайные силы, то наиболее вероятным, соответствующим максимуму энтропии, состоянием будет наиболее беспорядочное распределение направлений стрелок (рис.8). Если же, однако, стрелки намагничены, то наиболее вероятным, соответствующим максимуму энтропии, состоянием будет уже упорядоченное состояние,когда стрелки направлены параллельно друг другу (рис.9).

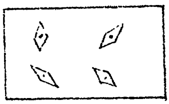

Рис.8

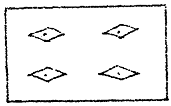

Рис.9

На молекулярном уровне электромагнитные силы взаимодействия приводят к появлению высокоупорядоченного кристаллического состояния вещества, когда молекулы (или ионы) располагаются правильными рядами, на упорядоченных расстояниях друг от друга.

Аналогично силы химического взаимодействия упорядочивают атомы, строя из них молекулы, где каждый атом занимает определенное место среди остальных.

Внимательно вглядываясь в природу, мы обнаружим в ней как силы, нарушающие порядок, так и силы, устанавливающие его. Вот как высказывается по этому поводу проф.А.И.Китайгородский в своей известной книге "Порядок и беспорядок в мире атомов": "До тех пор, пока в игру не вмешиваются силы, действующие на частицы, наиболее вероятным распределение молекул является беспорядок, как в отношении расположения, так и в отношении скоростей. Действие сил направлено к увеличению порядка. Если атомы (молекулы) находятся в тепловом движении и на них действуют силы, то наиболее вероятным распределением частиц уже не явится беспорядок".

В теоретической физике доказывается общее утверждение, что плотность атомов на единицу объема при наличии сил, имеющих потенциал, будет пропорциональная выражению $e^{-\frac{п.э.}{kT}}$, где $п.э.$ – выражение для потенциальной энергии атомов в функции координат("Фейнмановские лекции по физике", том 4, изд."Мир" 1965 г.). В частности, для атомов, находящихся в поле силы тяжести (атмосфера) плотность атомов n будет равна

$$n = n_0 e^{-\frac{mgh}{kT}},$$

где h – высота, а n_0 – плотность при $h = 0$. Таким образом, плотность атомов в единице объема экспоненциально убывает с увеличением высоты.

Рассмотрим теперь общий случай большого количества молекул, между которыми действуют силы, зависящие от расстояния между молекулами. Силу взаимодействия можно рассматривать как производную от потенциальной функции $V(r_{ij})$, где r_{ij} – расстояние между i-й и j-й молекулами. Типичный вид потенциальной функции $V(r)$ показан на рис.10. Если $r > r_0$, то при сближении молекул энергия уменьшается и молекулы притягиваются. Если же молекулы сходятся еще ближе, то перевес получают силы отталкивания.

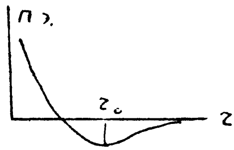

Рис.10

Полная потенциальная энергия равна сумме всех парных энергий и вероятность того, что молекулы образуют конфигурацию, характеризуемую заданными комбинациями расстояний r_{ij} пропорциональна выражению

$$e^{-\frac{1}{kT}\sum_{ij}^{n} V(r_{ij})}.$$

Если температура очень высока и $kT \gg V$, то показатели экспонент почти всюду малы, и вероятность найти молекулу в том или ином месте почти не зависит от ее расстояния до других молекул. Если же температура не велика, то экспонента максимальна при r близком к r_0, т.е. молекулы предпочитают быть на определенных расстояниях друг от друга. Таким образом, при наличии сил взаимодействия атомы сближаются, образуя упорядоченные структуры жидкостей, твердых тел и т.п. Наиболее вероятным, соответствующим максимуму энтропии, состоянием оказывается состояние известной упорядоченности (зависящей, разумеется, от функции $V(r)$

Только для частного случая идеального газа, молекулы которого взаимодействуют между собой как идеально-упругие шары, и потенциальную функцию $V(r)$ которого можно представить в виде кривой, показанной на рис.11 (где $r_ш$ – радиус шара), все конфигурации для любых $z > z_\omega$ равновероятны, и поэтому состояние с наибольшей энтропией (и вероятностью) будет наиболее неупорядоченным состоянием (вероятность найти молекулу в любой точке пространства не зависит от координат точки). Реальные вещества приближаются к состоянию идеального газа лишь при высоких температурах, когда средняя кинетическая энергия молекулы (она порядка kT) значительно превосходит потенциальную энергию сил взаимодействия, устанавливающих порядок.

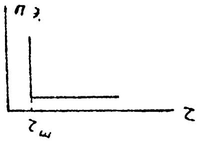

Рис.II

Таким образом, мы убеждаемся, что распространенное представление об энтропии как о мере неупорядоченности в действительности является ошибочным и не имеет под собой серьезной основы.

Каким же образом возникло и широко распространилось это ошибочное представление? На это трудно ответить с полной определенностью. Если мы обратимся к наиболее полным и авторитетным курсам физики (Л.Д.Ландау и Е.М.Лифшиц "Теоретическая физика" том 5, статистическая физика, издат. "Наука" I964 г., С.Э.Фриш, А.В.Тиморева, "Курс общей физики" том I, Физматгиз, I962 г.), то в них мы вообще не найдем трактовки энтропии как меры беспорядка. Действительно, в подробных курсах физики проводится термодинамический анализ систем, в которых действуют различные силы и делается ясным, что энтропия отнюдь не может выступать в качестве всеобщей меры неупорядоченности.

Однако в менее полных учебниках физики утверждения об энтропии как о мере неупорядоченности можно встретить. Это связано с тем, что в кратких учебниках энтропия рассматривается обычно на простейшем примере идеального газа, между молекулами которого не действует иных сил, кроме короткодействующих сил отталкивания. Понятие же энтропии, как понятие абстрактное, оказывается трудным для усвоения и поэтому возникает соблазн облегчить его понимание наглядным отождествлением энтропии с беспорядком, тем более что в рассматриваемом круге примеров (идеальный газ) это отождествление не ведет к противоречиям. Оговорка же о том, что отождествление это условно и справедливо лишь для ограниченного круга природных явлений, обычно опускается. Из этих учебников ошибочное отождествление энтропии с мерой неупорядоченности проникает в популярные книги и получает широчайшее распространение. Надо отметить, правда, что не только авторы популярных книг, но и некоторые крупные физики (например, Л.Бриллюэн) связывает энтропию с неупорядоченностью, но, как уже отмечалось, в наиболее авторитетных руководствах по физике трактовка энтропии как меры неупорядоченности отвергается, поскольку такая трактовка не основана на фактах и ведет ко многим противоречиям. Приведем в подтверждение выдержку из одного наиболее авторитетных физических журналов - "Успехи физических наук" (том 93, 1967 г. выпуск 4, статья С.Г.Суворова "К 50-летию со дня смерти Мариана Смолуховского"); "Классическая физика исходила из того, что состояние с максимальной энтропией - это газ, равномерно распределенный в пространстве. Теперь стало ясно, что это абстракция, вытекающая из пренебрежения взаимодействием между частицами. В космических масштабах благодаря тяготению происходит превращение более или менее однородного газа в галактики, звезды и т.д. Это возникновение структуры и отдельных тел является естественным процессом, сопровождается увеличением энтропии и находится в согласии с термодинамикой. Однако гравитационное воздействие является не единственной формой взаимодействия. Поэтому в определенных условиях возникают и другие структурные формы (молекулы, клетки и т.п.) Жупел "тепловой смерти" как превращения всего сущего в аморфное, бесструктурное образование является антинаучным, ошибочным по вполне конкретным физическим причинам".

Понимание энтропии как меры неупорядоченности распространено среди биологов* и оно приводит их к целому ряду противоречий. Действительно, рассматривая процессы жизнедеятельности различных организмов, легко убедиться, что эти процессы приводят к повышению упорядоченности системы. Иногда это объясняется тем, что живой организм не является изолированной системой, он является системой открытой, способной обмениваться с окружающей средой энергией и веществом, и повышение упорядоченности в самом организме неизбежно сопряжено с уменьшением ее во внешней среде, так что суммарная неупорядоченность изолированной системы - организм плюс внешняя среда - всегда увеличивается. Однако такое объяснение не является удовлетворительным: упорядоченность вследствие жизнедеятельности может возрастать и в изолированных системах. Действительно, рассмотрим колонию микроорганизмов, посаженную в изолированный сосуд с простой неупорядоченной смесью молекул питательных веществ и минеральных солей. Жизнедеятельность даже самой небольшой в начальный момент времени колонии приведет к тому, что неупорядоченная смесь молекул питательных веществ быстро превратится в упорядоченную - бактерии построят из молекул питательных веществ свои тела, заполнив вскоре ими почти весь сосуд (образуются и отходы - углекислота и вода - но их немного), а клетки бактерий - это весьма сложно организованные и упорядоченные молекулярные образования. Такое превращение неупорядоченности в упорядоченность, если рассматривать энтропию как меру неупорядоченности, будет, естественно, противоречить второму началу термодинамики. Можно, правда, возразить, что находящиеся в изолированном сосуде бактерии рано или позжно обречены на гибель, но и это не снимает противоречия, ибо даже временное повышение упорядоченности вследствие жизнедеятельности бактерий в сосуде, далеко выходит за рамки малой статистической флюктуации, допускаемой вторым началом. Примеры повышения упорядоченности вследствие деятельности живых организмов не единичны. Они постоянно встречаются в биологии. Сочетаясь с традиционной трактовкой энтропии как меры неупорядоченности, эти примеры выступают теперь как указание на кажущееся противоречие между законами управляющих жизнедеятельностью, и вторым началом термодинамики.

Так, в уже цитированной книге М.В.Волькенштейна "Молекулы и жизнь", мы читаем (с.30): "Здание, предоставленное самому себе, постепенно разрушается. Разумная деятельность живого организма - человека - направлена на преодоление второго начала. Человек строит здания и поддерживает их в порядке, он не увеличивает, а уменьшает энтропию". Но если признается наличие противоречия между жизненными явлениями и вторым началом термодинамики, то, очевидно, это противоречие нужно как-то объяснить. Как же оно объясняется? Одно из таких объяснений мы вотречаем у Н.Винера в его известной книге "Кибернетика" (изд."Советское радио", 1958, с.74-80). Н.Винер считает, что энзимы клетки, являющиеся катализаторами реакций, протекающих в клетке, можно рассматривать как своего рода "максвелловские демоны", которые, как говорит Винер, "уменьшают энтропию, но, может быть, не путем разделения быстрых и медленных частиц, а каким-нибудь другим эквивалентным способом". Таким образом, согласно Н.Винеру, специфические элементы живого организма – энзимы – могут на некоторое время замедлить неумолимый рост энтропии и неупорядоченности, но это время неминуемо является коротким и все жизненные явления предстают перед нами лишь как случайный недолговечный островок в огромном мире, где царствуют энтропия и хаос. "Стабильное состояние энзима наступапает, когда он перерастает действовать, а стабильное состояние живого организма наступает, когда он умирает" – пишет Н.Винер и продолжает: "Все катализаторы в конце концов отравляются, ибо они изменяют лишь скорость реакций, но не меняют истинного равновесия".

Аналогичных взглядов придерживался и Э.Кольман ("Философские проблемы кибернетики" изд. социально-экономической литературы, 1961): "С более общей, физической, точки зрения машина и живой организм сходны, потому что они представляют собой островки в океане возрастающей энтропии во всех макро-процессах на обитаемой нами части вселенной; это островки, где энтропия убывает, ибо накапливается информация".

Наиболее далеко идущие выводы из "противоречия" между процессами жизнедеятельности и вторым началом термодинамики делает К.С.Тринчер в своей выдержавшей два издания книге "Биология и информация (элементы биологической термодинамики") изд."Наука", 1964 г. (первое издание) и 1965 (второе издание).

К.С.Тринчер считает: "В рамках термодинамики необратимых стационарных процессов, как и в рамках классической термодинамики и химической кинетики невозможно объяснить прогрессивную эволюцию живой материи возникновение и сохранение организма в хаотической среде, развитие и размножение организма при стационарных условиях окружающей его закрытой системы, и, наконец, неограниченную морфо-физиологическую эволюцию всей совокупности организмов и в их онто- и филогенезе".

К.С.Тринчер мотивирует это утверждение тем, что второе начало термодинамики означает неизбежность прогрессирующей со временем потери структурности у любой машины", причем термин "машина" понимается в самом широком смысле, как совокупность элементов и частей любой степени сложности, взаимодействие между которыми может происходить любым образом, совместимым только с законами физики. К.С.Тринчер считает, что любая "машина", даже при столь широком определении, в хаотической среде обречена на гибель из-за неизбежной со временем потери своей структуры. "Работа машины - пишет К.С.Тринчер - это одновременно причина ее гибели. Состояние максимальной энтропии в изолированной системе означает не только превращение всей энергии системы в тепло, но и переход всех машин, т.е. всех программно движущихся структурированных образований в хаотически движущиеся неструктурированные вещества. Нет такой машины, которая работала бы, не увеличивая при этом энтропии за счет беспрерывной потери своей структуры. А так как "живые" машины - живые организмы - обладают противоположными свойствами, то К.С.Тринчер делает свой основной вывод: живые организмы не подчиняются обычным законам физики (и, в частности, - второму закону термодинамики). В них господствуют свои законы, отличные от законов неживой природы. К.С.Тринчер считает,что "объяснить возникновение, существование и эволюцию живых систем в рамках современных физических теорий, по-видимому, невозможно", и что "биологический обмен веществ противоречит второму начала термодинамики и не может быть материально смоделирован". Далее К.С.Тринчер утверждает, что "более 4-х миллиардов лет тому назад, когда Вселенная начинала свою нынешнюю фазу развития - фазу расширения - возникли (приблизительно в одно и то же время) две материальные сущности: живая и неживая материя и они развиваются с тех пор каждая по своим физическим законам".

Эти утверждения К.С.Тринчера противоречат основному научному принципу - принципу всеобщности законов физики. Законы эти справедливы для всех атомов и молекул, и не могут измениться от того, что атомы и молекулы вошли в состав живого тела.

Биологическая форма движения - как более высшая форма движения по сравнению с физической - может подчиняться новым, дополнительным законам и закономерностям, не сводимым к законам физики, но эти биологические закономерности ни в коей мере не могут нарушать физических законов - ни закона сохранения энергии, ни второго начала термодинамики.

Таким образом, мы убеждаемся, что представления об энтропии как о мере беспорядка приводят к весьма определенным следствиям: либо мы должны признать, следуя Н.Винеру и Э.Кольману, что все проявления жизни на Земле и все дела человеческих рук являются лишь случайным, недолговечным, не закономерным островком, неминуемо обреченным на уничтожение, либо же мы должны, следуя К.С.Тринчера, считать жизненные явления противоречащим законам физики и допускать, что молекулы кислорода, азота и т.п., войдя в состав живого организма, могут уже не подчиняться физическим законам.

Разумеется, если бы эти выводы основывались на фактах, то мы были бы обязаны принять их, несмотря на всю их странность и противоречие нашим представлениям. Однако в действительности в этом нет необходимости. Выводы Н.Винера, Э.Кольмана, К.С.Тринчера основаны на не обоснованной трактовке энтропии как меры неупорядоченности и являются ошибочными.

Второе начало термодинамики в действительности вовсе не запрещает существования в изолированных системах процессов, ведущих к увеличению упорядоченности (с одновременным увеличением энтропии). Процессы усложнения структуры, повышения упорядоченности могут протекать самопроизвольно, без какого-либо внешнего воздействия и достаточно часто (хотя и реже, чем противоположные им процессы) наблюдаются на практике.

Мы убеждаемся, что отказ от необоснованной трактовки энтропии как меры неупорядоченности позволяет иметь более реалистический взгляд на процессы жизнедеятельности животных и человека, на их место в общем ходе эволюции мироздания. Нам нет теперь необходимости признавать, что живые организмы или построенные человеком машины являются случайным, не закономерным, изолированным островком в окружающем океане косной материи. Второе начало термодинамики действительно устанавливает направленность протекающих в природе процессов – в макроскопических масштабах могут протекать только процессы, идущие с увеличением энтропии – но увеличение энтропии может сопровождаться как уменьшением, так и увеличением упорядоченности, и процессы увеличения упорядоченности (будь то рост и размножение сложно устроенных живых организмов или ширящееся распространение изделий человеческих рук) являются столь же закономерными процессами как и противоположные им процессы распада структуры, уменьшения упорядоченности.

Понимание второго начала термодинамики как закона о неминуемом росте неупорядоченности и энтропии как меры неупорядоченности подкупает своей простотой, кажущейся наглядностью и очевидностью. Сперва охватывает радостное чувство понимания: да, все ясно, второе начало объясняет господствующую в природе тенденцию к неупорядоченности, в ходе взаимодействий неизбежно разрушаются структура и порядок, если в одной части системы упорядоченность возросла, то в других частях она неизбежна (и более значительно) уменьшилась. Таков закон природы. Память быстро выдает примеры, подтверждающие этот "закон" и наступает довольство: "да, я понял".

При более глубоком анализе сомнения возникают вновь: ведь второе начало термодинамики - это универсальный закон. Он не может иметь исключений (во всяком случае в макропроцессах). А явления жизни? А процессы кристаллизации? А роль дальнодействующих сил? Действительно ли теоретическая физика рассматривает энтропию как меру неупорядоченности, или это только популярный, наглядный образ? В ходе анализа ответов на эти вопросы возникает новый уровень понимания, несколько менее наглядный, но более правильный. Второе начало термодинамики выступает теперь как закон перехода природных систем от менее вероятных состояний к более вероятным, а будут эти более вероятные состояния более упорядоченными, или менее упорядоченными чем исходные - это уже зависит от конкретной системы, от того, какие физические законы в этой системе действуют, какие силы преобладают. Да, действительно, очень часто более вероятное состояние оказывается менее упорядоченным. Это объясняет, почему так часто встречаются примеры, когда рост энтропии сопровождается возрастанием неупорядоченным, но существуют и примеры противоположные. Все зависит от конкретной системы, от характера сил, в ней действующих.

Новый уровень понимания, не теряя наглядность, позволяет более правильно ориентироваться в окружающем мире, причем не только в отвлеченно-теоретических вопросах, но и в делах чисто практических. Возьмем, например, такой вопрос, как защита природы, охрана ее от выбросов и отходов промышленности и жизнедеятельности человека. Если признать, что энтропия - это мера неупорядоченности, то тогда, если быть последовательными, надо согласиться с тем, что любая деятельность человека, увеличивая энтропию, увеличивает и хаос на Земле и защита природы может быть только пассивной. Человек может уменьшить наиболее вредоносные виды своей деятельности, может перенести ее неизбежные плоды (возрастающую неупорядоченность) в более отдаленные от себя районы, - например - используя канализацию, или строя высокие дымовые трубы - но его деятельность всегда увеличивает беспорядок в окружающей природе и тем шире объем его деятельности на Земле (а такое расширение неизбежно), тем больше вносимая неупорядоченность. К такому пессимистическому приводит - если быть последовательным - представление об энтропии как о мере неупорядоченности.

Если же второе начало термодинамики понимать как закон перехода системы от менее вероятных состояний к более вероятным, то становится оправданным более оптимистичный взгляд на отношения человека и природы, более активный подход к ее защите. Да, любая деятельность человека увеличивает энтропию. Но мы можем построить свою деятельность так, чтобы это увеличение энтропии сопровождалось увеличением порядка (и конкретно - порядка выгодного и удобного нам), причем не только в нашем ближайшем окружении, но и на всей Земле. Это не всегда легко сделать, но в принципе это возможно и на практике уже (к сожалению - не везде) делается. Хороший пример - разработки полезных ископаемых, например, - бурого угля в открытых карьерах. При таких разработках приходится перемещать большие массы грунта и часто после завершения добычи ископаемого на месте карьера остается нагромождение перевороченной земли, на которой ничего не растет. Бесплодную землю развевает ветер и на месте бывших разработок на долгие годы остается безотрадный хаотический ландшафт, который иронически называют "лунным пейзажем". Однако такая порча природы не является неизбежной. Если перед началом разработки карьера снять верхний слой плодородной почвы, сложить ее отдельно, а потом, после завершения работ по добыче угля, вновь разравнять ее сверху, то в этом случае более вероятным оказывается процесс быстрого восстановления растительности на насыпанном слое почвы, причем растет она даже лучше, чем на участках, не затронутых рукою человека. Такое грамотное ведение добычи уже введено на ряде карьеров (к сожалению не на всех).

Деятельность человека неизбежно увеличивает энтропию Земли, но совсем не обязательно должна увеличивать неупорядоченность. Отказ от упрощенного понимания энтропии как меры неупорядоченности позволяет выработать более правильный взгляд на природные процессы и, в частности, на проблему охраны природы.

Остается последний вопрос - представление об энтропии как о неупорядоченности наглядно, просто, в настоящее время - уже привычно. Нужно ли от него отказываться безусловно и считать его только ошибкой? Поскольку существует очень большое количество систем, в которых наиболее вероятное состояние является вместе с тем и состоянием наибольшей неупорядоченности, а во всех этих системах рост энтропии сопровождается и ростом неупорядоченности, то упрощенное представление об энтропии во многих случаях оказывается правильным и полезным. Им пользовался и сам Л.Больцман. Нужно лишь не забывать, что абсолютизация представления об энтропии как о мере неупорядоченности, распространение его "на все случаи жизни", в том числе и на системы, где наиболее вероятное состояние, не есть состояние наибольшего беспорядка может привести к ошибкам.

Следует отметить также, что критические замечания, сделанные автором по поводу некоторых толкований второго начала термодинамики и понятий упорядоченности и энтропии, не должны создавать у читателя представления, что принципиальные научные вопросы, связанные со вторым началом, к настоящему времени являются спорными или неясными. Общепризнано, что второе начало термодинамики является всеобщим законом природы, и что любая изолированная система с течением времени стремится к увеличению энтропии. В этих принципиальных вопросах разногласий нет. Они начинаются на другом уровне, на уровне истолкования, на уровне наглядного представления законов, существующих в природе. Означает ли рост энтропии увеличение неупорядоченности, обязательно ли изолированная система стремится к состоянию наибольшего беспорядка - вот тут уже возникают иной раз и разногласия.

Это связано, конечно, с тем, что энтропия - понятие научное, точно определенное, ее можно вычислить и измерить, а там, где царствует число и мера, разногласия довольно быстро изживаются. Совсем другое дело – такие понятия, как беспорядок, упорядоченность. Это понятия житейские, они трудно поддаются точному определению и количественной оценке, поэтому возможность разногласий здесь гораздо больше.

§ 3. СПРАВЕДЛИВ ЛИ ПРИНЦИП МИНИМУМА ПРОИЗВОДСТВА ЭНТРОПИИ В СТАЦИОНАРНОМ РЕЖИМЕ?

Важность второго начала термодинамики заключается прежде всего в его всеобщности. Даже если изолированная система является очень сложной, и мы не знаем законов взаимодействия ее элементов, мы можем утверждать, что с течением времени она придет к равновесному состоянию, к состоянию с максимальной энтропией и нулевым производством энтропии, а это уже довольно много говорит нам о свойствах систем.

Однако второе начало термодинамики относится лишь к изолированным системам, к системам, которые не обмениваются с окружающей средой ни энергией, ни веществом. В природе не менее часто встречаются открытые системы, в которых происходит беспрерывный обмен веществом или энергией с окружающей средой. Если параметры окружающей среды неизменны, то в открытой системе с течением времени устанавливается стационарный режим, основные параметры которого постоянны и не зависят от времени.

Для исследования сложных систем очень важно выделить свойства и характеристики стационарного режима, отличающие его от режимов нестационарных. И вот в 1945 г. бельгийский ученый Илья Романович Пригожин выдвинул интересную гипотезу об открытых системах (заметим, что И.Р.Пригожин родился в 1917 г. в Москве, однако в 1920 г. он был увезен родителями за границу и основные работы выполнил в Бельгии; в 1977 г. он был удостоен Нобелевской премии). Гипотеза (или принцип) И.Р.Пригожина заключается в том, что в стационарном режиме любой открытой системы производство энтропии минимально по сравнению с режимами нестационарными, отвечающими тем же условиям на границе между системой и окружающей средой. Вот одна из последних по времени формулировок этого принципа, данная самим И.Р.Пригожиным в лекции при вручении ему Нобелевской премии: "Теорема о минимуме производства энтропии ... утверждает, что производство энтропии системой, находящейся в стационарном достаточно близком к равновесному состоянию, минимально". И далее там же: " ... когда граничные условия позволяют системам достичь термодинамического равновесия, система останавливается в состоянии "минимальной диссипации" [10].

Принцип Пригожина очень удобен при исследовании сложных систем, он быстро получил весьма широкую популярность и стал считаться принципом доказанным, на который можно опираться. Этому способствовал и высокий личный авторитет И.Р.Пригожина - лауреата Нобелевской премии. Принцип Пригожина вошел в число основных принципов неравновесной термодинамики и стал широко использоваться при решении различных задач в области физики, биологии, сложных систем управления (см.работы [5, 6, 11] и др.).

Однако при применении принципа Пригожина к решению практических задач наряду с правильными результатами еще в 1963-64 гг. стали появляться и результаты ошибочные. Это обстоятельство побудило проанализировать - действительно ли принцип Пригожина носит всеобщий характер? Анализ затруднялся тем, что в работах самого Пригожина и его последователей [4, 5] принцип минимума производства энтропии выдвигался сразу для весьма сложных систем и поэтому в его обоснование выдвигались общие наводящие соображения, а не точные расчеты, которые для сложных систем крайне затруднительны. Необходимо было найти такую систему, в которой производство энтропии поддается точному и неопровержимому расчету.

В качестве такой системы рассмотрим однородный твердый брус длиной L и сечением S (в дальнейшем положим для простоты $L=1$ и $S=1$). Боковые стенки бруса теплоизолированы, на левом торце бруса поддерживается температура "a" градусов (в дальнейшем для простоты $a=1$), а на правом торце - $1+m$. Поэтому тепло распространяется по брусу справа налево. Брус расположен вдоль оси Ox. Процесс распространения тепла в брусе будет, как известно, описываться линейным уравнением теплопроводности:

$$\frac{\partial T}{\partial t} = a\,\frac{\partial^2 T}{\partial x^2}\,, \tag{18}$$

где T - температура бруса, являющаяся функцией времени t и координаты x , причем $0 \le x \le 1$, a – коэффициент теплопроводности, постоянная величина. Таким образом, рассматриваемая нами система линейна. Заданы граничные условия: $T(x=0; t) = 1$; $T(x=1; t) = 1+m$. В стационарном режиме температура не зависит от времени, $\frac{\partial T}{\partial t} = 0$, и тогда из уравнения (I8) и граничных условий находим, что в стационарном режиме $T_{ст} = 1+mx$, т.е. температура зависит от координаты x линейно.

Подсчитаем теперь производство энтропии. Выделим внутри бруса элементарный слой толщиной dx. Через слой будет проходить поток тепла $a\frac{\partial T}{\partial x}$, а поскольку поток тепла входит в слой при одной температуре, а выходит при другой, то производство энтропии в элементарном слое согласно формуле (I) равно

$$d\Pi = a\frac{\partial T}{\partial x}\left(\frac{1}{T} - \frac{1}{T+\frac{\partial T}{\partial x}dx}\right) = a \cdot \frac{1}{T^2}\left(\frac{dT}{dx}\right)^2, \tag{I9}$$

а во всем брусе производство энтропии

$$\Pi = \int_0^1 d\Pi = a\int_0^1 \frac{1}{T^2}\left(\frac{dT}{dx}\right)^2 dx . \tag{20}$$

Вычисляя интеграл (20) для стационарного режима, когда $T = 1+mx$, а производная $\frac{dT}{dx} = m$, находим, что в стационарном режиме производство энтропии равно

$$\Pi_{ст} = a\,\frac{m^2}{1+m} . \tag{2I}$$

Однако это значение не является минимальным. Обычными приемами вариационного исчисления нетрудно установить (вычисления приведены в [9]), что минимум производства энтропии достигается при нестационарном, экспоненциальном распределении температуры. Заданным граничным условиям удовлетворяет распределение:

$$T_{min} = (1+m)^x . \tag{22}$$

Подставив функцию (22) в интеграл (20), вычислим минимальное производство энтропии:

$$\Pi_{min} = a\left[\ln(1+m)\right]^2 . \tag{23}$$

Нетрудно проверить, что для любых $m > 0$ будет $П_{ст} > П_{мин}$, т.е. производство энтропии в стационарном режиме минимума не достигает. Так, если $m = 1,72$, то $П_{мин} = a$ в то время как $П_{ст} = 1,08\,a$. Этот пример был опубликован в [8]. По сути дела после этого на принципе Пригожина можно было поставить точку. Ведь если обнаружен хотя бы один опровергающий пример (контрпример), то это означает, что принцип Пригожина - не всеобщий, и прежде чем использовать его в каком-либо конкретном случае для исследования стационарного режима какой-либо конкретной системы, нужно сперва доказать, что именно для этой системы принцип Пригожина справедлив. А поскольку доказать это, конечно, сложнее, чем непосредственно исследовать характеристики стационарного режима, то ясно, что при наличии контрпримера никакой эвристической силой принцип Пригожина обладать не может. Если же отдельного доказательства справедливости принципа Пригожина для каждой рассматриваемой системы не делать, то можно придти к ошибочным заключениям.

Однако, поскольку принцип Пригожина к 1966 г. был уже признанным, основополагающим принципом неравновесной термодинамики, то против контрпримера, опубликованного в [8], выставлялись возражения.

В одном из возражений указывалось, что разница в производстве энтропии при стационарном и при экспоненциальном распределении температуры является (для малых m) величиной высшего порядка малости, т.е. принцип Пригожина верен хотя бы как приближенный. Действительно, разложив в ряд по степеням m отношение

$$\frac{П_{ст} - П_{мин}}{П_{ст}} = \frac{1}{12} m^2 - \frac{1}{4} m^3 + \ldots \qquad (24)$$

находим, что оно имеет второй порядок малости по отношению к m. Но если мы вычислим разность температур в стационарном режиме и при экспоненциальном распределении температуры (22), соответствующем минимуму производства энтропии, то найдем, что эта разность также мала. Она достигает максимума вблизи $x = 0,5$ и имеет тот же самый порядок малости по отношению к m. Так, при $x = 0,5$ имеем

$$\frac{T_{ст} - T_{мин}}{T_{ст}} = \frac{1}{8} m^2 - \frac{1}{8} m^3 + \ldots \qquad (25)$$

Кроме того, принцип Пригожина - утверждение не количественное, а качественное. Он утверждает, что в любом нестационарном режиме производство энтропии больше, а рассмотренный пример показывает, что существуют нестационарные режимы, в которых производство энтропии меньше, чем в стационарном.

Были возражения, утверждавшие, что принцип Пригожина справедлив только для линейных систем и только при малых отклонениях от равновесного состояния (для бруса равновесное состояние соответствует $m = 0$). Однако линейность уравнения (18) и разложение (24) показывают, что принцип Пригожина может быть несправедлив и для линейных систем, причем при сколь угодно малых отклонениях от равновесного режима.

В одном из возражений утверждалось, что в стационарном режиме достигает минимума не производство энтропии, а скорость диссипации свободной энергии. Эта скорость, как показывает неравенство (5), пропорциональна произведению производства энтропии на абсолютную температуру T. Для рассмотренного нами примера с брусом скорость диссипации свободной энергии равна

$$J_{дисс} = a \int_0^1 \frac{1}{T}\left(\frac{dT}{dx}\right)^2 dx \qquad (26)$$

и тоже, как нетрудно проверить, не достигает минимума в стационарном режиме. Минимума в стационарном режиме достигает интеграл

$$J = \int_0^1 \left(\frac{dT}{dx}\right)^2 dx \,, \qquad (27)$$

не тождественный интегралу (26).

Наиболее интересен и оригинален был отклик самого И.Р.Пригожина на контрпример, опубликованный в [8]. Этот отклик был опубликован в книге [4]. Признав, в отличие от других, прямо, что даже при рассмотрении простого процесса теплопереноса в твердом теле могут существовать примеры, когда в стационарных режимах производство энтропии не минимально, Пригожин в книге [4] утверждал, что в этих случаях будет достигать минимума производство так называемой "взвешенной энтропии" $L = S \cdot T^2$, равной произведению настоящей энтропии на квадрат абсолютной температуры T, и что этот факт якобы спасает универсальность принципа минимума производства энтропии. Сам по себе факт верен: при теплопереносе в брусе "взвешенная энтропия" действительно достигает минимума в стационарном режиме. Однако спасает ли это универсальность принципа Пригожина? Нет, не спасает. "Взвешенная энтропия" имеет другую размерность и совсем другой физический смысл, чем энтропия настоящая. Нетрудно показать, что при переходе тепла от горячего тела к холодному настоящая энтропия S как и следовало ожидать, возрастает, а "взвешенная энтропия" убывает.

Кроме того, остается неясным для каких систем в стационарном режиме достигается минимум производства настоящей энтропии, и для каких – "взвешенной".

Таким образом, возражение Пригожина остроумно, но несостоятельно. Однако оно было составлено так тонко, что и после его опубликования в [4], большинство исследователей продолжали верить в универсальность принципа Пригожина, многие продолжают опираться на него при анализе конкретных систем и их стационарных режимов, а ведь это может приводить и к ошибкам!

Так, например, в работе [II] принцип Пригожина был использован для отыскания распределения тока в массиве металла с учетом поверхностного эффекта. Использование принципа Пригожина действительно очень облегчило анализ распределения тока. Но можем ли мы быть уверены, что результат, приведенный в [II] верен? Ведь для уверенности в его правильности мы должны доказать, (а доказать это труднее, чем непосредственно вычислить распределение тока без использования принципа Пригожина, что для исследуемой в [II] системы – массива металла, в которой проникает внешнее переменное магнитное поле – в стационарном режиме производство энтропии действительно минимально, причем минимально именно производство настоящей энтропии S , а не "взвешенной" энтропии $L = S\,T^2$ Пока не доказана минимальность производства настоящей энтропии S для системы, рассматриваемой в [II], мы не можем быть уверены, что результат, приведенный в [II], действительно верен.

Материал настоящего параграфа показывает, с какой осторожностью следует в научной работе опираться даже на казалось бы признанные положения. И это заставляет нас с особым вниманием отнестись к любимому девизу К.Маркса "подвергай все сомнению".

§ 4. ЭНТРОПИЯ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

В предыдущих параграфах анализировалось понятие энтропии в физике. В настоящем параграфе мы рассмотрим совсем другую энтропию – энтропию распределения вероятностей, используемую в теории информации.

Представим себе опыт, который может иметь несколько возможных исходов с различными вероятностями. Пусть, например, наугад вынимается шар из урны, содержащей 5 белых, 3 черных и 2 красных шара. Опыт имеет три исхода (появление белого, черного или красного шара) с вероятностями $p_1 = 0{,}5$, $p_2 = 0{,}3$ и $p_3 = 0{,}2$. Если же урна содержит 8 белых и 2 черных шара, то вероятности исходов будут $p_1 = 0{,}8$ и $p_2 = 0{,}2$. Таким образом, для каждого опыта появляется свое распределение вероятностей $p_1; p_2; \cdot\ p_i$ (отметим, что всегда $\Sigma p_i = 1$). Энтропией распределения вероятностей будем называть сумму произведений вероятностей p_i на их логарифмы $\log p_i$, взятую с обратным знаком, т.е.

$$H = - \sum_{i=1}^{n} p_i \log p_i \tag{28}$$

Знак минус в формуле (28) взят потому, что поскольку все вероятности $p_i < 1$, то логарифмы их отрицательны и поэтому сумма $\Sigma p_i \log p_i$ тоже отрицательна. Знак минус обеспечивает тем самым положительность величины H. Логарифмы в формуле (28) могут быть взяты по любому основанию. Обычно пользуются логарифмами по основанию два, двоичными логарифмами. Если логарифмы взять по другому основанию m, величина H окажется умноженной на постоянное число $a = \log_2 m$ (модуль перехода).

Для рассмотренного нами примера (выбора шара из урны, где 5 белых, 3 черных и 2 красных шаров) энтропия распределения вероятностей $H = 1{,}48$. Аналогично при извлечении шара из урны, где 8 белых и 2 черных шара, имеем $H = 0{,}72$, а если в урне 5 белых и 5 черных шаров, то $H = 1$.

Таким образом, H является числовой безразмерной характеристикой, характеризующей распределение вероятностей исходов опыта. Если число исходов фиксировано, то величина H тем больше, чем ближе друг к другу вероятности различных исходов. Если все p_i

равны между собой, $P_i = \frac{1}{n}$, то H достигает (как легко доказать) максимального значения

$$H_{макс} = \log_2 n. \qquad (29)$$

При переменном числе исходов n величина H тем больше, чем больше число исходов. Мы убеждаемся, таким образом, что энтропия распределения вероятностей характеризует неопределенность исходов опыта, трудность предсказания его результата. Действительно, если в урне находится I0 шаров и все они разного цвета, то предсказать, какого цвета будет вытянутый шар намного труднее, чем если бы в урне было 5 белых и 5 черных шаров (в первом случае $H = \log_2 10$ = = 3,02, а во втором H = I).

Если же в урне 9 белых шаров и I черный, то предсказать результат еще проще, в этом случае H = 0,47).

Заметим, что величина H не может, естественно, характеризовать всех сторон случайного явления или процесса. Она характеризует только вероятности исходов, оставляя в стороне разнообразие возможных значений самих исходов.

Рассмотрим для примера лотерею, участник которой с вероятностью 0,09 выигрывает I0 руб. и с вероятностью 0,0I - 50 руб. (т.е. с вероятностью 0,9 не выигрывает ничего). Выигрыш участника лотереи является здесь случайной величиной, колеблющейся от 0 до 50 руб. (математическое ожидание выигрыша M = 0,09хI0+0,0Iх50 = = I,4 руб.

Рассмотрим теперь вторую лотерею, участник которой с вероятностью 0,09 выигрывает 5 руб. и с вероятностью 0,0I - 95 руб. Легко видеть, что выигрыш участника второй лотереи (при том же математическом ожидании M = 0,09х5+0,0Iх95 = I,4 руб. будет более неопределенной, более "разбросанной" случайной величиной, в то же время энтропия распределения вероятностей различных выигрышей для обеих лотерей остается одной и той же (в обоих случаях H = $= -0{,}9\log_2 0{,}9 - 0{,}09\log_2 0{,}09 - 0{,}01\log_2 0{,}01$ = 0,47). В данном случае степень неопределенности величины выигрыша лучше характеризует известная в теории вероятностей еще с XIX века характеристика - дисперсия, т.е. суммы произведений квадратов отклонений значений случайной величины от математического ожидания на их вероятности. Для первой лотереи дисперсия выигрыша $D_1 = 0{,}09\ 8{,}6^2 + 0{,}01 \cdot 48{,}6^2$ =

= 29,266, для второй лотереи дисперсия $D_2 = 0{,}09 \cdot 3{,}6^2 + 0{,}01 \cdot 93{,}6^2 =$ = 88,776.

Таким образом, энтропия распределения вероятностей не является исчерпывающей характеристикой случайной величины и не может поэтому быть всеобщей "мерой неопределенности" случайного процесса или явления. Она характеризует трудность предсказания появления того или иного исхода случайного явления, но не характеризует неопределенности, разброса значений самих исходов. Несмотря на это функция H во многих случаях является очень удобной, особенно в теории связи, где важную роль играет именно степень непредсказуемости появления того или иного сигнала в передаваемом сообщении. Введенная впервые К.Шенноном в 1948 г. функция H быстро получила самое широкое применение.

Остановимся теперь на происхождении названия этой функции, которую чаще всего называют просто энтропией, вместо полного и правильного названия энтропия распределения вероятностей. Выбор названия предопределило сходство формулы (28), определяющей функцию H , с ранее приводимой (в § I) формулой (I3), выражающей термодинамическую энтропию для частного случая идеального газа, на молекулы которого не действуют никакие силы, кроме сил отталкивания между молекулами. Имея в виду это сходство, К.Шеннон предложил в 1948 г. назвать введенную им функцию H , определенную равенством (28) энтропией множества вероятностей $P_1, P_2, \ldots P_n$ [I3]. В дальнейшем многочисленные последователи и популяризаторы К.Шеннона стали отождествлять функцию H с энтропией в физике, что привело к большой путанице и многим ложным выводам.

Отметим, что редактор первого перевода основной работы К.Шеннона на русский язык, профессор Н.А.Железнов писал в 1953 г. в специальном примечании к переводу: "автор (К.Шеннон) вводит термин "энтропия" на основании чисто внешнего сходства выражения для введенной им величины H с выражением для энтропии в общепринятом значении. Поскольку с понятием энтропии в статистической физике связано вполне определенное физическое содержание, то во терминологии автора, в дальнейшем слово "энтропия" поставлено в кавычки. Следует иметь в виду, что в данном случае слово "энтропия" (в кавычках) есть не более, чем краткое название величины

$$H = \sum_{i=1}^{n} p_i \log p_i , \qquad (30)$$

где p_i - вероятность появления некоторого события i " [12]. К сожалению, все предостережения остались гласом вопиющего в пустыне. Кавычки, введенные Н.А.Железновым, постепенно исчезли, смешение функции H с энтропией физики стало свершившимся фактом и не могло не привести ко многим ошибочным представлениям. Поэтому во избежание недоразумений функцию H следует называть ее полным названием - "энтропия распределения вероятностей", имея в виду, что краткий термин "энтропия" имеет в физике другой смысл. Энтропия физики является размерной величиной, пропорциональной логарифму вероятности состояния системы и характеризующей направленность процессов, происходящих в ней, в то время как энтропия распределения вероятностей теории информации - безразмерная характеристика вероятностей исходов опыта, производимого над системой и характеризующая трудность предсказания этих исходов. Общей характеристикой или мерой неопределенности она быть не может - это показывает приведенный нами пример с лотереей.

§ 5. ИНФОРМАЦИЯ

Если термин "энтропия" является искусственным словом, введенным в употребление Р.К.Клаузиусом в середине XIX века специально для научных целей, то центральное понятие теории информации - понятие "информация" - имеет в основе обиходное слово "информация" имеющее смысл "знания", "сообщения", "новости".

Однако наука не может основываться на обиходном значении слова. Научное понятие - а особенно такое фундаментальное понятие как "информация" - должно иметь четкое научное определение. В действительности же теория информации не имеет однозначного и общепринятого определения своего основного понятия. До сих пор - хотя, разумеется, это достаточно странно - в самом определении информации, а тем самым и в понимании ее сущности, наблюдаются значительные разногласия. Правда, некоторые авторы считают, что без определения можно обойтись. Вот что пишет, в частности, Ф.П.Тарасенко (Введение в курс теории информации, Изд. Томского ун-та, 1963): "Как это ни парадоксально звучит, но для развития теории информации в ее современном виде вообще не требуется определения понятия информации как таковой; необходимым и достаточным для построения теории является понятие количества информации. Поэтому употребление терминов "информация" и "количество информации" как синонимов не вызывает недоразумений". Далее Ф.П.Тарасенко поясняет: "это не должно казаться странным: такое положение характерно и для ряда других количественных теорий. Например, для изложения механики нужны лишь количественные характеристики движения, но не требуется анализа существа самого движения". Однако в действительности отсутствие определения информации не являлось препятствием лишь на первом этапе развития теории, когда область применения теории ограничивалась почти исключительно техникой связи. По мере расширения круга приложений, когда методы теории информации проникли в биологию, в физику, в теорию автоматического управления и даже в область искусства и литературы, нечеткость в определении основного понятия не могла не привести к целому ряду затруднений, которые интерпретировались иногда даже как "кризис" всей теории информации.

Вот как высказывается по этому поводу В.С.Флейшман ("Конструктивные методы оптимального кодирования для каналов с шумами" изд. АН СССР, 1963) "возникшая благодаря гениальной интуиции К.Шеннона, исходившего из физических предпосылок измерения величины информации - этого "флогистона" кибернетики - теория информации в настоящее время испытывает кризис неадекватности физических представлений и своего аппарата ... живой интерес к теории информации и интенсивное развитие ее в первые годы после опубликования работы Шеннона в 1948 г. сменились в последние годы заметным охлаждением интереса к этому направлению".

Поэтому следует с особым вниманием отнестись к вопросу определения понятия "информация". Различными авторами в разное время было дано большое количество разнородных определений. Эти определения можно разбить на следующие три основные группы:

1. Описательные определения.

2. Определения, связывающие информацию с отражением.

3. Определения, связывающие информацию с энтропией, понимаемой как мера неупорядоченности.

Разберем эти определения по группам.

Описательные определения.

I. Терминология, рекомендуемая АН СССР: "информация" - это сведения, являющиеся объектом хранения, передачи, преобразования.

2. А.А.Красовский, Г.С.Поспелов (Основы автоматики и технической кибернетики. Г.Э.И. 1962): "информация - это особая совокупность сведений, первичным источником которых является опыт".

3. Н.Винер (Кибернетика. Советское радио, 1958): "информация - это информация, а не материя и не энергия".

4. Н.М.Амосов (Моделирование информации и программ в сложных системах// Вопросы философии. 1963. № 12): "информация - это содержание воздействия, его величина, изменение в пространстве и времени, используемое как средство связи сложных систем".

5. В.М.Глушков (Мышление и кибернетика// Вопросы философии. 1963. № I): "информация есть мера неоднородности материи и энергии в пространстве, мера изменений, которыми сопровождаются все протекающие в мире процессы".

Эти определения нельзя признать удовлетворительными. Действительно, информация - это фундаментальное понятие, определяя ее через "сведения", мы должны определить - что такое сами "сведения", а этого нельзя сделать не прибегая вновь к понятию информации - получается логический круг.

Определение Н.М.Амосова узко: если исходить из него, то получается, что понятие информации приложимо лишь к связи сложных систем.

Определение В.М.Глушкова, напротив, необъятно широко, и в этом его слабость. Оно не выделяет специфики информации и информационных процессов, в то время как хорошее определение как это требовал еще К.Вейерштрасс, "должно охватывать сущность предмета, содержать, хотя бы в неявном виде, возможность построения полной научной теории".

Другая группа определений определяет информацию через философское понятие отражения.

Наиболее просто к этому вопросу подходит И.Б.Новик (Негэнтропия и количество информации// Вопросы философии. 1962. № 6); согласно ему, информация есть просто "мера отражения".

Несколько более осторожно подходит к этому же вопросу Ф.П.Тарасенко (К определению понятия "информация" в кибернетике//Вопросы философии. 1963. № 4): "понятие информации в кибернетике родственно с понятием отражения в диалектическом материализме. Они являются разными абстракциями одного и того же свойства материи".

Согласно А.М.Коршунову и В.В.Мантатову (Гносеологический анализ понятия "информация"// Методологические проблемы современной науки. МГУ. 1964) информация есть "активная форма отражения".

Все эти определения неудовлетворительны потому, что "отражение" - это философское понятие, а выводы и понятия философии должны основываться на выводах и понятиях конкретных наук, являться их обобщением, но не наоборот. То есть можно определять отражение через информацию (если предварительно само понятие "информация" хорошо определено и всесторонне понятно), но не следует определять информацию через отражение.

Третью группу определений информации составляют определения, связывающие информацию с энтропией (причем в этих определениях отождествляется, как правило, энтропия физики с энтропией распределения вероятностей и обе "энтропии" понимаются как "мера неопределенности" системы). Традиция связывать информацию с энтропией идет от Н.Винера, который в своей известной книге "Кибернетика" утверждал, в частности, что "понятие количества информации совершенно естественно связывается с классическим понятием статистической механики, с понятием энтропии. Как количество информации в системе есть мера организованности системы, точно так же энтропия есть мера неорганизованности системы. Одно равно другому, взятому с обратным знаком".

Согласно А.Г.Ивахненко ("Техническая кибернетика", Киев, 1959): "энтропия понимается как мера вероятности состояния, мера дезориентации элементов, мера хаоса. Информация понимается как мера особенности состояния, мера упорядочения, порядка". Аналогичных взглядов придерживается и Б.С.Украинцев ("Информация и отражение", Вопросы философии, 1963, № 2): "В кибернетике количество информации в системе принимается за меру организованности системы, а энтропия системы - за меру ее неорганизованности".

Все эти определения несостоятельны уже потому, что основываются на неверном представлении об энтропии как о мере неупорядоченности системы. Ошибочность такого представления была рассмотрена ранее. Не является обоснованным и представление об информации как о мере порядка. К определению понятия информации надо идти другим путем.

§ 6. ОПРЕДЕЛЕНИЕ ИНФОРМАЦИИ

Для того чтобы подойти к определению информации, надо внимательно проанализировать специфические черты процессов, которые мы считаем информационными и отличающие их от процессов, не носящих информационного характера.

Основным объектом изучения в науке долгое время являлись системы, находящиеся в устойчивых, стабильных состояниях. Причиной изменений в таких системах является воздействие внешних сил.

Для детального описания процессов, происходящих при взаимодействиях материальных систем был введен ряд характеристик, помогающих описывать как механизм различного рода взаимодействий, так и его результаты. Такими характеристиками являлись понятия силы, количества движения, момента количества движения (удобны для описания механических воздействий), количество тепла, температуры (удобны для тепловых взаимодействий) и т.п.

Наиболее универсальной характеристикой взаимодействия, полезной для очень широкого класса взаимодействий, является энергия. Важность понятия "энергия", введенного в науку в XIX веке, общеизвестна.

Однако существуют такие классы взаимодействий, где результат воздействия одной системы на другую практически не зависит от энергии (а также от количества движения, момента количества движения, температуры и других, подобных этим, характеристик воздействия).

Примером такого рода взаимодействий как раз и служат те взаимодействия, с которыми приходится сталкиваться в технике связи, при передаче сигналов. Радиоприемник принимает радиосигнал (и выдает на выходе сигнал звуковой) почти независимо от энергии радиосигнала. Эта энергия может изменяться в миллионы раз, но до тех пор пока она выше чувствительности приемника, он будет реагировать на сигналы разной интенсивности практически одинаково.

Это связано с тем, что приемник имеет свои внутренние источники энергии (например, батареи), а роль энергии внешнего радиосигнала - только вспомогательная.

Внешний радиосигнал только распоряжается этой внутренней энергией приемника, направляет ее по различным путям (в безбатарейном (детекторном) приемнике в акустическую энергию преобразуется энергия приходящего радиосигнала, но и здесь существенна не энергия сигнала, а его модуляция).

Если мы проанализируем внимательно основные черты всех тех взаимодействий, которые имеют место в технике связи при передаче различных сигналов и сообщений, то мы увидим, что все их объединяет общая черта - недостаточность для описания результатов взаимодействия введенных ранее характеристик - энергии, количества движения, температуры и т.п. Настоятельно требуется новая характеристика. Такой характеристикой и является информация.

Рассмотрим два примера приемный телеграфный буквопечатающий аппарат. На вход аппарата поступают электрические импульсы разной длительности ("точки" и "тире"), на выходе - на бумажной ленте отпечатываются буквы и слова. Результаты воздействия входных импульсов в очень широких пределах не зависит от их энергии, напряжения, силы тока и т.п. Он зависит от комбинаций приходящих импульсов, от того - в каком порядке и в какой последовательности приходят они на аппарат. Совершенно естественно, что требуется новая мера, отличная от энергии и ей подобных мер, которая характеризовала бы разнообразие комбинаций приходящих импульсов - такой мерой и будет являться информация.

Таким образом, мы подходим к определению понятия "информации": информация - это мера особого класса взаимодействий материальных систем.

Чем характерен этот класс? Чем он отличается от распространенных энергетических взаимодействий? Перечислим основные различия.

I. Результат взаимодействия не находится в непрерывной зависимости от энергии воздействия. Часто вообще нельзя уловить зависимость результата взаимодействия от энергии.

2. Взаимодействие несимметрично, в нем можно выделить систему воздействующую и систему, воспринимающую воздействие, причем

часто система, оказавшая воздействие, не претерпевает изменений в результате взаимодействия. В то же время энергетические взаимодействия (вследствие законов сохранения энергии, количества движения и т.п.) всегда симметричны - сколько энергии отдала одна система, участвующая во взаимодействии, столько получила вторая.

3. Система, испытывающая воздействие, находится в метастабильном состоянии. Это означает, что ее внутренние процессы могут протекать различными путями, приводить к различным результатам в зависимости от внешних воздействий малой интенсивности ("сигналов"), причем различие результатов взаимодействия зависит не от энергии сигналов, а от других характеристик. Так например, в буквопечатающем телеграфном аппарате печатаемая буква зависит не от энергии поступающих импульсов, а от их порядка во времени; поведение человека, слушающего собеседника, зависит не от громкости речи последнего, а от ее смысла - т.е. от порядка следования произносимых звуков (фонем), в радиоприемнике процессы на выходе (т.е. воспроизводимая им речь или музыка) определяются не энергией радиосигнала (лишь бы она была больше определенного уровня), а его модуляцией.

Взаимодействия, обладающие этими чертами, выделим в особый класс и назовем сигнальными взаимодействиями.

Теперь получаем возможность дать точное определение понятия "информация": информация есть мера сигнальных взаимодействий, характеризующая разнообразие возможных результатов взаимодействия.

Что нового влечет за собой это определение?

Во-первых, оно включает процессы передачи информации в технике связи и обмена сообщениями между людьми в общий и широкий мир сигнальных взаимодействий, имеющих место в самых различных системах как живой, так и неживой природы. Тем самым подводится база под расширение круга приложений понятия "информация", применение его в физике, биологии, в области искусства и литературы и т.п.

Во-вторых, определение подчеркивает, что информация есть характеристика, мера взаимодействия, по крайней мере двух материальных систем, и поэтому количество информации зависит, естественно, от свойств обоих взаимодействующих систем. Этот вывод важен потому, что многие парадоксы и неясности в теории информации проистекают от забвения этого обстоятельства. Так, рассматривается иногда совершенно нелепый, не имеющий смысла вопрос "сколько информации содержится в книге, телеграфном сообщении и т.п." не оговаривая - для кого, для какой взаимодействующей системы. Такая путаница происходит от того, что в технике связи удобно рассматривать изолированно свойства сигнала, свойства канала связи и связисты фактически рассматривают (не оговаривая этого явно; отсюда и путаница) идеализированную приемную систему.

В-третьих, определение позволяет классифицировать различные формы информации на разных ступенях развития живого. Простейшими формами информации являются формы чистого взаимодействия, управления, когда одна система непосредственно направляет по другому пути процессы в другой системе. Так, для мышцы сердца адреналин в крови является как сигналом, свидетельствующим о необходимости мобилизации сил организма, так и непосредственно (ферментативно) воздействует на сердце, увеличивая частоту и силу сокращений.

Для пчел в улье химическое вещество (феромон) слизываемое с тела матки при кормлении ее и перераспределение потом среди всех рабочих пчел несет, с одной стороны, информацию о наличии матки в улье, с другой стороны - непосредственно перестраивает жизненные процессы рабочих пчел (кормление молоди) так, чтобы новые матки в улье не выводились. Только при отсутствии в улье маточного феромона пчелы строят маточные ячейки и выкармливают маток.

В этих простейших случаях особенно четко выступает существо информации как меры взаимодействия.

На более высокой стадии развития живых организмов (и особенно - в человеческом обществе) информация выступает в более сложной форме.

У высших животных сигналы из внешнего мира могут служить уже не только побуждением к непосредственному действию, но и могут запоминаться, запасаться впрок, для использования в дальнейшем. Когда животное прислушивается, принюхивается к обстановке (или человек читает письмо, сообщение), то происходит взаимодействие между окружающей средой и клетками мозга, взаимодействие, скрытое от наших глаз и внешне сразу не проявляющееся. Поэтому в своей развитой форме информация выглядит как некоторые "знания", "сообщения", а ее сущность как меры взаимодействия - затушевывается и не проступает явно.

Однако на самом деле значение информации хранящейся в мозгу, заключается в конечном счете именно в том, чтобы изменять, направлять действия животного (или человека). Так что и здесь информация в конечном счете является мерой взаимодействия. Более того, только по объективно проявляющимся результатам взаимодействия можно вообще судить о наличии или отсутствии информации. Ведь на основании чего можно сделать заключение о том, что план города, например, заключает в себе информацию о расположении улиц города? Только на основании того, что действия человека, ознакомившегося с планом, будут во многом аналогичны действиям того, кто ознакомился с городом "в натуре".

Информационные цепочки

Анализируя процессы передачи и приема информации можно убедиться, что в ряде случаев сигналы, несущие информацию, могут проходить через ряд преобразующих звеньев. Так, человеческая речь одного из собеседников (акустические колебания) может быть микрофоном превращена в колебания тока, идущего по телефонному кабелю, а затем вновь превращена в акустические колебания приемной трубкой телефона и услышана другим собеседником. Еще более длинную цепочку преобразований может претерпеть информация, идущая к читателю газеты от ее корреспондента (речь корреспондента (очевидца событий) - микрофон - магнитофонная лента - секретарь редакции - пишущая машинка - наборная машина - газетный оттиск - читатель).

Вообще говоря, можно каждое среднее звено информационной цепочки рассматривать и как передатчик и как приемник информации (по отношению к предыдущему и последующему звеньям). Однако более правильным будет рассматривать в качестве приемника информации только последнее звено цепочки. Процессы, происходящие в средних звеньях рациональнее рассматривать просто как процессы преобразования энергии приходящих сигналов в другую форму (акустических колебаний в электрические и т.п.). Процессы, происходящие в средних звеньях цепочки, могут и не обладать характерными признаками информационных процессов. Надо всегда искать самые важные звенья – начало и конец цепочки и приемник информации.

§ 7. О ВЫЧИСЛЕНИИ КОЛИЧЕСТВА ИНФОРМАЦИИ

В настоящем параграфе мы рассмотрим методику вычисления количества информации, предложенную К.Шенноном. Будем называть ее классической методикой, ибо работу К.Шеннона, опубликованную в 1948 г. можно с полным основанием считать классической.

Рассмотрение ограничим простейшим случаем канала связи без шумов. Напомним, что если понятие "информации" в классической теории не определялось строго, то количество информации во всяком случае определялось и вычислялось вполне четко и однозначно. Анализ методики вычисления количества информации поможет уяснить характер зависимости между информацией и энтропией распределения вероятностей передаваемых символов и выявит как неоспоримые достоинства классической теории, так и ее ограниченность.

Для вычисления количества информации существуют две формулы – формула Хартли и формула Шеннона, причем формула Шеннона является существенным обобщением и улучшением формулы Хартли; по сути дела с момента опубликования формулы Шеннона (1948 г.) и ведет свое начало современная теория информации. В чем суть формул Хартли и Шеннона?

К своей формуле Хартли пришел от анализа количественной стороны телеграфной связи. Если телеграмма передается с помощью двух символов – точек и тире – то число различных телеграмм длиною в n символов будет $N = 2^n$ Число N возрастает экспоненциально с ростом n ; так, при n = 10, N = 1024, а при n = 20, N = 1048576 или в 1024 раза больше. Экспоненциальная зависимость N от n не позволяет выбрать само число N в качестве меры. Действительно, передача телеграммы длиной в 20 знаков требует, очевидно, не в 1024 раза, а в два раза больше труда и времени, чем передача телеграммы длиной 10 знаков. Поэтому в качестве меры количества информации Хартли предложил логарифм N, т.е.

$$\underline{I} = \log N,$$

причем логарифмы можно брать по любому основанию. Обычно их берут по основанию два (двоичные логарифмы), и тогда количество информации в простейшем сообщении, для которого возможны два значения (точка или тире, да или нет), будет равно единице. Эту единицу информации называют битом.

Определение Хартли пригодно и для сообщений, состоящих из последовательности не двух, а любого числа символов.

Так, если сообщение передается буквенным алфавитом, с числом букв 32, то число различных слов, которые можно составить из четырех букв, будет равно $32^4 = (2^5)^4 = 2^{20} = 1048576$, а количество информации, содержащейся в одном четырехбуквенном слове, будет равно

$$I = \log_2 2^{20} = 20 \text{ бит}.$$

В общем случае, когда алфавит состоит из m символов, то число возможных сообщений длиной в n символов будет равно $N = m^n$ а количество информации в одном сообщении

$$I = \log m^n = n \log m. \tag{31}$$

Это количество можно определить и по-другому. Действительно, если число всех возможных сообщений N, то вероятность того, что будет получено именно данное сообщение, равна $1/N$. Можно количество информации определить как логарифм (с обратным знаком) вероятности сообщения, т.е.

$$I = -\log \frac{1}{N}. \tag{32}$$

Это определение тождественно с предыдущим, поскольку $-\log \frac{1}{N} = \log N$. Точно так же для сообщения длиной n символов из алфавита с количеством символов m вероятность появления в сообщении именно данного символа равна $1/m$ (если все символы равновероятны), вероятность последовательности из n символов равна $1/m^n$ а количество информации

$$I = -\log \frac{1}{m^n} = -n \log \frac{1}{m}, \tag{33}$$

что совпадает с предыдущей формулой, поскольку $-n \log \frac{1}{m} = n \log m$

Определение количества информации, данное Хартли, пригодно в том случае, если все символы, встречающиеся в передаче, встречаются в ней одинаково часто, имеют одинаковую вероятность появления. Так, если речь идет о передаче буквенного текста, где символами являются буквы, то определение Хартли пригодно, если в среднем любая буква в достаточно длинном сообщении будет встречаться одинаковое число раз. Однако для всех реальных языков это условие не выполняется (в русском языке буквы О, е, а встречаются много чаще, чем ф, э, щ; в английском буквы s, t, e – чаще, чем q, x, z). Поэтому в 1948 г. Клод Шеннон предложил новое определение количества информации, существенно расширившее определение Хартли.

Представим себе сообщение, состоящее из достаточно длинной последовательности символов, имеющих разные вероятности (например, сообщение на русском языке, составленное из букв русского алфавита). Пусть всего имеется m символов (в случае русского алфавита – 32) и вероятность каждого символа – p_i (где i = 1, 2, ..., n). Вероятность сигнала, составленного из n символов, равна

$$p = p_1 \cdot p_2 \cdot \ldots \cdot p_n \tag{34}$$

(n множителей).

Если каждый i -й символ встречается в сигнале n_i раз, то

$$p = p_1^{n_1} \cdot p_2^{n_2} \cdot \ldots \cdot p_m^{n_m} \tag{35}$$

(m множителей; естественно, $m \leq n$).
Но при достаточно длинном сигнале любой i -й символ встретися число раз $n_i = p_i \cdot n$ т.е. n_i равно общему числу символов n, умноженному на вероятность i -го символа p_i. Следовательно

$$p = p_1^{np_1} \cdot p_2^{np_2} \cdot \ldots \cdot p_m^{np_m}, \tag{36}$$

и количество информации I будет равно

$$I = -\log p = -n \sum_{i=1}^{m} p_i \log p_i \,. \tag{37}$$

Это соотношение носит название формулы Шеннона. Она является обобщением формулы Хартли. Действительно, если все символы равновероятны, т.е. все $p_i = 1/m$ то формулу Шеннона переходит в формулу Хартли, так как

$$-n \sum_{i=1}^{m} \frac{1}{m} \log \frac{1}{m} = -n \log \frac{1}{m}$$

Если сравнить формулу Шеннона для вычисления количества информации и формулу для энтропии идеального газа, то между ними обнаруживается известное сходство. Это сходство и побудило Клода Шеннона назвать величину $H = -\sum_{i=1}^{m} p_i \log p_i$ "энтропией множества вероятностей $p_1; p_2, \dots p_m$".

Эта же формула явилась также поводом к трактовке информации как отрицательной энтропии, "негэнтропии". Действительно, в результате приема сообщения неопределенность возможного сигнала на приемной стороне (численно равная как раз энтропии распределения вероятностей H) снимается, уничтожается. Принятый сигнал является уже вполне определенным сигналом, неопределенность его равна нулю. Так как снятие неопределенности, обращение в нуль энтропии распределения вероятностей H произошли в результате приема информации, то отсюда и возникло представление об информации как об отрицательной энтропии множества вероятностей передаваемых символов, а затем, вследствие смещения величины с энтропией в физике, распространилось уже представление об информации просто как об отрицательной энтропии, негэнтропии, "мере упорядочения, порядка", "мере организованности" и т.п. (см. § 5). Мы убеждаемся теперь, что такое представление нельзя признать правильным уже хотя бы потому, что энтропия в физике не является "мерой неупорядоченности", а энтропия распределения вероятностей не тождественна физической энтропии.

Формула Шеннона относится, по самому своему выводу, к информации, содержащейся в достаточно длинных сообщениях. С.Голдман (Теория информации. ИЛ, 1959) дал определение количества информации в единичном сообщении, исходя из количества неопределенности, которое оно устраняет. Пусть до получения информации вероятность некоторого события равна p_i , а после сообщения о том, что оно произошло, неопределенность исчезла (вероятность обратилась в единицу). Тогда количество информации по С.Голдману равно логарифму отношения вероятностей

$$I = \log \frac{1}{p_i} = -\log p_i .$$

Чем невероятнее сообщение, тем больше информации, по С.Голдману, оно содержит.

Для последовательности сообщений определение С.Голдмана дает те же результаты, что и определение Шеннона. Действительно, если передается последовательность сообщений и вероятность появления i-го сообщения равна p_i, а количество информации в нем, по С.Голдману, равно $-\log p_i$, то в "n" сообщениях количество информации I равно

$$I = -\sum_{i=1}^{n} n p_i \log p_i = -n \sum_{i=1}^{n} p_i \log p_i ,$$

что совпадает с формулой Шеннона.

Если говорить о сообщениях, записанных на реальных языках (английском, русском и т.д.), то количество информации, в каком либо отрезке текста, вычисленное по формуле Шеннона с учетом вероятностей появления отдельных букв, сочетаний из двух, трех и т.п. букв примерно в два раза меньше, чем вычисляемое по формуле Хартли (для английского языка эта оценка установлена К.Шенноном, для русского – Ю.Добрушиным).

Теперь, ознакомившись с определениями количества информации, предложенными Голдманом, Шенноном и Хартли, можно уяснить, в чем заключаются как достоинства классической теории информации, так и ее недостаточность.

Прежде всего отметим, что классическая теория предполагает наличие идеального наблюдателя, воспринимающего сигналы. В определении Хартли предполагается наблюдатель, для которого каждая комбинация кодовых символов различима и имеет смысл, определение Шеннона предполагает наличие "несколько менее идеального" наблюдателя, для которого различима и имеет смысл каждая комбинация символов, удовлетворяющая вероятностным законам появления символов и их комбинаций, характерным для данного языка.

Реальные получатели информации отнюдь не обладают свойствами идеальных приемников – это предопределяет неизбежную ограниченность классической теории, формулы которой дают, как правило, очень сильно завышенные результаты для количества информации.

Рассмотрим для примера, отца с нетерпением ожидающего из родильного дома известия о том, кто родится - мальчик или девочка. Такой отец является системой с двумя состояниями и сообщение "родился мальчик" несет ему один бит информации. В то же время, согласно определению Хартли, в четырнадцатибуквенном сообщении "родился мальчик" содержится количество информации

$$I = 14 \log_2 32 = 70 \text{ бит},$$

а согласно определению Шеннона - примерно в два раза меньше, т.е. примерно 35 бит. Таким образом, формула Шеннона несколько лучше формулы Хартли, но и она дает сильно завышенные результаты.

Теперь рассмотрим, что нового может принести подход, опирающийся на определение информации как меры особого класса взаимодействий.

Прежде всего отметим, что аналогия с другими классами взаимодействий подсказывает, что может быть несколько видов информации и количество информации для разных видов может вычисляться по различным формулам.

Так, например, для механического взаимодействия существуют различные меры - количество движения (mv), момент количества движения, механическая энергия ($\frac{m}{2} v^2$). Каждая из этих мер наилучшим образом отражает и описывает какую-то одну сторону механических взаимодействий и все они взаимно дополняют друг друга.

Естественно предположить, что и информационные взаимодействия имеют несколько мер.

Действительно, если рассмотреть подробнее процессы передачи и приема информации, то можно подметить, что в разных случаях целевое назначение передаваемой информации может быть совершенно различным. В одним случаях информация непосредственно служит какой-то четко определенной цели (информация: "где это находится?" "Как пройти туда" и т.п.). В других случаях информация выступает как чистая передача знания, без ясно выраженной (по крайней мере в ближайшем будущем) функции практического использования (например - сообщение о научном открытии). Наконец, информация может выступать и в форме, отвлеченной от реального получателя - в такой форме она выступает в технике связи, где работа телеграфного кабеля, например, оценивается лишь по статистическим свойствам передаваемых сигналов (полоса частот и т.п.) и не зависит от того, передается по нему осмысленная речь, или бесформенный набор звуков.

Именно для вычисления количества этого частного вида информации и пригодны формулы Хартли и Шеннона. Использование их для количественной оценки других видов информации затруднительно и может приводить ко многим недоразумениям и парадоксам. Так,например, согласно формуле Шеннона (37) количество информации всегда, для любых p_i является положительной величиной. В то же время сообщение, например, о местоположении предмета, может быть как правильным, так и ложным, дезоринтирующим. Естественно считать, что в первом случае информация сообщения будет положительной, во втором - отрицательной ("дезинформация"). Классическая теория информации не позволяет охватить эти случаи, дать количественную меру "дезинформации".

Ограниченность классической теории была давно подмечена; на это указывал, например, А.А.Харкевич в своей работе "О ценности информации" (сборник "Проблемы кибернетики",вып.4, 1960 г.). Работы М.М.Бонгарда ("Проблема узнавания", изд. "Наука", 1967, глава УП "Полезная информация") и Ю.А.Шрейдера ("Об одной семантической модели информации". Проблемы кибернетики, выпуск 13, 1965) также можно рассматривать как исследования, посвященные различным видам информации, отличным от информации, понимаемой в смысле К.Шеннона.

§ 8. РАЗЛИЧНЫЕ ВИДЫ ИНФОРМАЦИИ

Информация классической теории связи, информация, мерой количества которой является энтропия распределения вероятностей передаваемых символов, является лишь одним из видов, одной из форм информации. Ее можно назвать статистической информацией. Действительно, в теории связи содержание передаваемых сообщений, их возможное использование - не существенны. Для работы приемно-передающих устройств играют роль не смысл сигналов, а их статистические свойства, разнообразие сигналов, их непредсказуемость, избыточность и т.п. Поэтому в классической теории связи за меру количества информации взята энтропия распределения вероятностей появления сигналов. Для удобства оценки свойств источника информации и канала связи в теории связи отвлекаются от свойств приемника, отвлекаются от того, что передаваемые сигналы могут быть лишены для него смысла и значения. Неявно предполагается, что каждая комбинация кодовых символов является одинаково важной. Информация в классической теории связи является тем самым мерой взаимодействия между источником передаваемых символов и "идеальным приемником", для которого каждая комбинация символов различима и имеет равный смысл и значение. Несмотря на то что именно статистической информации посвящено подавляющее большинство работ в области теории информации, необходимо помнить, что статистическая информация классической теории связи является лишь одним из видов информации, далеко не исчерпывающем разнообразия различных возможных форм сигнального взаимодействия между материальными системами. Забвение этого обстоятельства приводит к целому ряду парадоксов и недоумений.

С другим видом, с другой формой информации мы сталкиваемся в тех случаях, когда принимаемые сигналы используются для какой-либо конкретной цели, для повышения эффективности некоторой операции. Эту форму информации можно назвать целевой информацией и оценивать количество ее по приращению целевой функции, достигаемому вследствие получения информации. На практике цели, для которых используется информация, а с ними и целевые функции могут быть самыми различными, поэтому целевая информация будет разделяться на ряд подвидов, по разному вычисляемых. Остановимся на одном из этих подвидов, исследованном подробно М.М.Бонгардом ("Проблема узнавания", глава УП). Там приведен следующий иллюстрирующий пример: "некто X хочет застать в учреждении, открытом с I0 до I8 часов сотрудника У, о котором известно, что он бывает там по два часа ежедневно (кроме воскресенья). Желая уточнить эти сведения, X обратился к знакомому, работающему в том же учреждении. Знакомый ответил, что У бывает на рабсте после I4 часов в два раза чаще, чем до I4. После этого X позвонил секратарю в учреждение. В ответ на свой вопрос он услышал: "На будущий месяц расписание еще не составлено, то товарищ У всегда принимает пять раз в неделю с I2 до

I4 и один раз в неделю с I4 до I6. Ой, простите, я ошиблась: один раз с I2 до I4 и пять раз с I4 до I6". Последний ответ соответствовал действительности.

Какую полезную информацию получил X по интересующему его вопросу:

I) из ответа знакомого?

2) из ответа секретаря?

3) получил бы из ответа секретаря, если бы она не исправила ошибки?

Очевидно, для определенного ответа на эти вопросы нужно указать, как будет измеряться неопределенность задачи. М.М.Бонгард предлагает следующее определение: пусть выдвинута гипотеза, что распределение вероятностей наличия товарища X на месте в определенные часы есть q_1; q_2; ... q_i где i – порядковый номер часа, а истинное распределение вероятностей равно p_1; p_2; ... p_i. Тогда неопределенность задачи равна

$$N = -\sum_{i=1}^{n} p_i \log q_i .$$

Это определение справедливо для любой гипотезы о распределении вероятностей q_1; q_2; ... q_i некоторого события.

Неопределенность будет минимальна в том случае, если выдвинутая гипотеза соответствует истине; тогда

$$N = H = -\sum_{i=1}^{n} p_i \log p_i ,$$

т.е. неопределенность равна энтропии распределения вероятностей H ; если $q_i \neq p_i$ то, как легко доказать $N > H$.

М.М.Бонгардом доказано, что если какая-либо информация о распределении вероятностей q_i отсутствует, то на первом этапе целесообразной всего полагать все q_i равными друг другу (первоначальная гипотеза).

Для рассматриваемой нами задачи вероятности пребывания товарища У в учреждении с I0 до I2 часов, с I2 до I4 и т.д. по различным гипотезам будут равны:

	Часы суток			
	10–12	12–14	14–16	16–18
Истинное распределение вероятностей	0	$\frac{1}{6}$	$\frac{5}{6}$	0
Первоначальная гипотеза	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
Гипотеза после сообщения знакомого	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$
Гипотеза после ответа секретаря	0	$\frac{1}{6}$	$\frac{5}{6}$	0
Гипотеза после ошибочного ответа секретаря	0	$\frac{5}{6}$	$\frac{6}{6}$	0

Если выражать неопределенность и информацию в битах, то начальная неопределенность (по первоначальной гипотезе) будет равна

$$N_0 = -\frac{1}{6}\log_2\frac{1}{4} - \frac{5}{6}\log_2\frac{1}{4} = 2 \text{ бита}.$$

Неопределенность, соответствующая гипотезе уточненной после ответа знакомого

$$N_1 = -\frac{1}{6}\log_2\frac{1}{6} - \frac{5}{6}\log_2\frac{1}{3} \approx 1{,}75 \text{ бита}.$$

Неопределенность после ответа секретаря (гипотеза соответствует истине)

$$N_2 = -\frac{1}{6}\log_2\frac{1}{6} - \frac{5}{6}\log_2\frac{1}{6} \approx 0{,}65 \text{ бита}$$

Если бы секретарь не исправила ошибки, то неопределенность была бы равна

$$N_3 = -\frac{1}{6}\log_2\frac{5}{6} - \frac{5}{6}\log_2\frac{1}{6} \approx 2{,}2 \text{ бита}$$

Таким образом, количество целевой информации, полученное от знакомого

$$I_{зн} = N_0 - N_1 = 0{,}25 \text{ бита}.$$

Ответ секретаря содержит

$$I_{секр} = N_1 - N_2 = 1{,}1 \text{ бита}.$$

Если же разговор с секретарем происходил бы до ответа знакомого, то в том же ответе секретаря содержалось бы

информации $I_{секр_1} = N_0 - N_1 = 1,35$ бита

Наконец, ошибочный ответ секретаря нес отрицательную информацию:

$$I_{секр.\,ош} = N_1 - N_3 = -0,45 \text{ бита},$$

или 0,45 бита дезинформации.
М.М.Бонгард называет введенную им меру "полезной информацией". Однако полезными могут быть различные виды информации. Поэтому правильнее пользоваться термином "целевая информация".

Рассмотренный пример сразу показывает основные отличия целевой информации от статистической:

I. Целевая информация может быть не только положительной, но и отрицательной величиной.

2. Она зависит от свойств приемника, в частности – от той информации, которой он уже располагал до прихода нового сообщения.

Очевидно, что для различных практических задач целевые функции могут быть разными и мера количества целевой информации, предложенная М.М.Бонгардом, не является исчерпывающей (хотя М.Бонгардом доказано, что она применима для достаточно широкого круга задач). Изучение целевой информации (в отличие от статистической) только начинается.

Однако и целевая информация, при всем разнообразии целевых функций, не исчерпывает разнообразия встречающихся на практике различных видов информационных процессов. В частности, встречаются случаи, когда получаемая информация совсем не предназначена для какой-либо определенной цели, а просто расширяет объем знаний, сведений о предмете, причем совершенно неизвестно, когда и как эта информация будет использована.

Очевидно, что в этом случае, в случае информации, используемой для накопления знаний (назовем ее смысловой информацией), оценка количества информации по любой целевой функции – непригодна (тем более непригодна и классическая оценка по энтропии распределения вероятностей передаваемых символов).

Для того чтобы найти пути построения количественной оценки смысловой информации, рассмотрим процессы, которые происходят при получении информации, не предназначенной к немедленному использованию. Информация "запасаемая впрок", служит, очевидно для того, чтобы подготовить условия для использования другой информации, которая будет получаться впоследствии, но которая сама по себе, без предварительного запаса знаний, не может быть использована.

Так, например, когда оператор противовоздушной обороны получает информацию о силуэтах самолетов противника, о признаках, позволяющих отличить одни типы самолетов от других и свои самолеты от вражеских, то эта информация (слово "информация" в данном случае эквивалентно более узкому термину "знания") необходима для того, чтобы в будущем, при получении сигналов (зрительных или радиолокационных) о подозрительных самолетах оператор мог осуществить правильную последовательность действий. Таким образом, суть смысловой информации, получаемых знаний, заключается в том, что на их основе вырабатывается и закрепляется в мозгу соответствие между будущими символами или образами на входе и необходимыми действиями. Высшие животные не могли бы существовать, если бы у них в мозгу не существовало достаточно богатого соответствия (т.е. как бы "смыслового словаря") между разнообразными сигналами ситуаций внешней среды (запах хищника, запах добычи и т.п.) и целесообразными в этих ситуациях действиями. Причем, если у насекомых, например, такое соответствие является, в основном, врожденным и поведение их формируется на базе врожденных инстинктов (определяется наследственной информацией), то у млекопитающих оно пополняется в течение жизни животного и получаемая животным информация от органов чувств носит во многом смысловой характер, идет на пополнение запаса соответствий в их памяти. У человека смысловая информация может идти на пополнение соответствий не только между внешними сигналами и действиями, но и между понятиями.

Таким образом, общей чертой смысловой информации является то, что она изменяет запас сведений, запас соответствий (между сигналами и действиями, между понятиями и т.п.) у получателя информации. Первоначальный (и меняющийся в ходе получения информации) запас соответствий можно представить себе как некоторый обобщенный словарь или справочник, который Ю.А.Шрейдер предложил называть "тезаурусом" (от греческого "тезаурос" - сокровище; в словарной практике тезаурусами называют словари, где даны не только значения слов, но и связи между ними). В качестве меры количества смысловой информации естественно взять изменение тезауруса приемника под действием поступившей информации. Некоторые соображения о методике вычисления количества смысловой информации для некоторых частных случаев приведены в уже упоминавшейся статье Ю.А.Шрейдера ("Об одной самантической модели информации"). Мы не будем пока останавливаться на количественной стороне вопроса, поскольку изучение смысловой информации находится еще в самой начальной стадии и методы расчета будут еще неоднократно совершенствоваться. Ограничимся качественными соображениями, которые, впрочем, сами по себе представляют интерес.

Заметим прежде всего, что количество смысловой информации зависит от приемника, от его тезауруса. Если тезаурус приемника слишком беден, количество информации в сообщении может оказаться вообще равным нулю. Так, если сообщение - это учебник по высшей математике для I-го курса ВУЗ'а, то трехлетний ребенок получит от него информацию практически равную нулю. Школьник старших классов извлечет уже кое-что. Максимальную информацию извлечет очевидно, студент того курса, для которого учебник предназначен. В предельном, идеальном случае полного понимания и восприимчивости количество смысловой информации (в битах), равно информации статистической, информации в смысле К.Шеннона. Любопытно, что по мере дальнейшего расширения тезауруса приемника воспринимаемая информация начинает уменьшаться. Уже у студента II курса при чтении учебника I курса изменение тезауруса будет меньше, а для человека, хорошо знающего высшую математику, учебник I курса может уже не нести никакой информации.

Остановимся на сходстве и различии между статистической, целевой и смысловой информациями.

Смысловая информация (также как и целевая) самым существенным образом зависит от свойств приемника. Одно и то же сообщение, один и тот же сигнал может для разных приемников нести самую различную информацию - это мера взаимодействия между (по крайней мере) двумя материальными системами, и если статистическая информация, на первый взгляд, не зависит от свойств приемника, то лишь потому, что в ее определение неявно входит идеальный приемник, для которого каждая комбинация кодовых символов различима и приводит к ответной реакции.

Для смысловой информации (в отличие от статистической) наличие предварительного запаса сведений может как увеличивать количество информации в воспринятом сообщении, так и уменьшать его. Мы уже упоминали, что студент I курса получит из учебника высшей математики больше информации, чем школьник, поскольку он подготовлен к восприятию ее. В классической же теории связи увеличение априорного запаса сведений может только уменьшить количество информации в принятых сигналах: чем больше мы знаем о системе, тем меньше ее непредсказуемость и тем меньшую информацию несет сообщение об ее состоянии. В отличие же от целевой информации, смысловая информация всегда неотрицательна и может быть только равной нулю.

Правильное различение различных видов информации помогает разобраться во многочисленных парадоксах и недоумениях при попытках применения понятия информации к различным явлениям.

Сколько информации содержится в сообщении "на Марсе есть жизнь"? Согласно классической теории связи - I бит (действительно неопределенность ситуации мала - есть жизнь - нет жизни). Однако эта оценка - I бит - возбуждает эмоциональный протест. Действительно, инстинктивно ясно, что в сообщении "есть жизнь на Марсе" суть не в статистической, а в смысловой информации, а преобразование тезауруса (во всяком случае достаточно развитого, чтобы понять всю важность сообщения) вызванное сообщением такого рода будет очень велико.

Рассматривая информацию, заключенную в отрывке художественной прозы для нормального читателя (а не для исследователя, подсчитывающего, например, процентное соотношение гласных и согласных букв), легко убедиться, что к этой информации следует подходить как к информации смысловой, а не как к статистической или целевой информации. Действительно, прочитанная художественная литература не передается никуда по каналам связи, не используется для конкретной цели, но преобразует внутренний мир, изменяет "тезаурус" читателя, причем (как правило, хотя возможны и исключения) изменение "тезауруса" будет тем больше, чем талантливее, чем художественнее произведение. Таким образом, смысловая информация действительно может служить для оценки художественности произведения (по крайней мере в принципе; мы отвлекаемся пока от трудности выработки правильной количественной оценки).

Что произойдет если к художественной прозе подойти с позиций теории информации, но понятие "информация" понимать только в узком смысле классической теории связи и оценивать через энтропию распределения вероятностей? Попытки такого подхода делались (А.М.Кондратов "Математика и поэзия" изд. "Знание" 1962, В.Зарецкий "Образ как информация" Вопросы литературы, 1963, № 2 и т.п.). Сравнивалась, например, энтропия распределения вероятностей появления различных букв, слов в художественной прозе и в "штампованной" газетной речи, с ее однообразным языком, где по началу фразы можно обычно предсказать ее конец. Поскольку непредсказуемость, энтропия распределения вероятностей художественной прозы выше, чем у газетной речи, делался вывод, что информация в отрывке художественной прозы выше, чем информация в равном по длине отрывке газетного текста. Это положение соответствует нашим интуитивным представлениям и оценкам и поэтому делался вывод о применимости понятий классической теории связи к литературным произведениям.

Однако если продолжить этот анализ, то поскольку наивысшей энтропией распределения вероятностей обладает бессмысленная, совершенно случайная последовательность букв, приходится признать, что наибольшую информацию (на единицу длины текста) несет в себе текст, лишенный всякого смысла. С точки зрения теории связи этот вывод верен, так как передавать по каналам связи бессмысленный текст действительно труднее, чем осмысленный, поскольку в нем труднее исправить случайные ошибки, но этот вывод показывает только малую плодотворность применения понятия статистической информации к анализу языка и художественной литературы.

Действительно, вычисляя количество информации по формуле К.Шеннона мы фактически вычисляем информацию взаимодействия между литературным текстом и "идеальным приемником", для которого любая последовательность букв (удовлетворяющая частотным характеристикам повторения их в языке), в том числе и "заумная" и совершенно бессмысленная последовательность равнозначны и приводят к равнозначным ответным реакциям.

Совершенно очевидно, что реальный читатель вовсе не является таким "идеальным приемником" (приближается к такому идеалу разве что гоголевский Петрушка) и поэтому концепции классической теории связи в области искусства и литературы не оказались плодотворными. В хорошо аргументированной статье "Сорок лет спустя, число и чувство меры в изучении поэзии" (Вопросы литературы. 1963. № 4) профессор Л.Тимофеев подверг анализу исследования начала шестидесятых годов по приложению математики и теории информации к проблемам литературы и поэзии и показал, что эти исследования, использующие статистическую теорию информации, не принесли почти ничего нового и ценного по сравнению с работами двадцатых годов, когда о теории информации еще и помину не было.

Почему появились такие работы? Вероятнее всего, причина в том, что молодой науке – а теория информации наука молодая – свойственно преувеличивать свои силы и возможности. Основатель теории информации Клод Шеннон еще в 1956 г. указывал на опасность такой переоценки. "Теория информации – писал он – как модный опьяняющий напиток кружит головы всем вокруг. Для тех, кто работает в области теории информации, такая широкая популярность несомненно приятна и стимулирует их работу, но такая популярность в то же время и настораживает. Сознавая, что теория информации (К.Шеннон говорит все время о статистической информации) является сильным средством решения проблем теории связи (и в этом отношении ее значение будет возрастать) нельзя забывать, что она не является панацеей для инженера связиста и тем более для представителей всех других специальностей. Очень редко удается открыть одновременно несколько тайн природы одним и тем же ключом. Здание нашего несколько искусственно созданного благополучия слишком легко может рухнуть, если только в один прекрасный день окажется, что при помощи нескольких магических слов, таких как информация, энтропия, избыточность ..., нельзя решить всех нерешенных проблем" [13].

В области искусства и литературы методы теории информации станут плодотворными тогда, когда будут разработаны хорошие способы оценки количества не только статистической, но и семантической, смысловой информации, методы оценки изменения "тезауруса" читателя (в том числе и его эмоционального "тезауруса") под действием произведений искусства.

На сегодняшний день таких методов еще нет.

Расширение понятия "информация", отказ от слишком узкой трактовки ее, признание того, что наряду со статистической информацией классической теории связи существуют и имеют не меньшее значения другие виды, другие формы информации безусловно позволит расширить круг приложений теории информации, позволит более правильно подойти к анализу различных видов информационных процессов.

Надо иметь в виду, что если уже сравнительно простое, механическое взаимодействие материальных систем имеет несколько мер, характеризующих различные стороны взаимодействия (количество движения mv , кинетическая энергия $\frac{m}{2}v^2$ и т.п.), то вполне естественно ожидать, что более сложный класс взаимодействий - класс сигнальных, информационных взаимодействий - будет иметь не меньшее разнообразие мер, принимаемых к различным формам взаимодействий внутри этого класса.

Исчерпывается ли разнообразие различных видов информации рассмотренными выше формами - статистической, целевой и смысловой информации? Очевидно, нет. Можно рассматривать как отдельную форму информации, например, командную, управляющую информацию, когда роль принятого сигнала (приказа) заключается в том, что он направляет по определенному пути процессы, которые до приема сигнала имели некоторое разнообразие возможных путей. Количество информации в приказе определяется, очевидно, разнообразием возможных действий у исполнителя (приемника информации) до получения приказа. Целесообразно ввести логарифмическую меру - т.е. если до получения приказа исполнитель мог совершить одно из n возможных действий (или комбинаций действий), то количество командной информации

$$I = \log_2 n \text{ битов.}$$

Таким образом, количество командной информации вычисляется по формуле, аналогичной формуле Хартли для статической информации. Однако, в отличие от статистической, количество командной информации не зависит от вероятностей реализации различных возможностей действий до получения команды, и зависит не только от характеристик сигнала, но и от свойств получателя информации.

Рассмотрим в качестве примера управление движением фигур на автоматизированной шахматной доске. Пусть командный сигнал является последовательностью нулей и единиц. Если мы управляем движением короля, то поскольку за один ход король может перейти (если он не стоит на краю доски и рядом нет мешающих фигур) максимум на одно из восьми соседних полей, информация в управляющем сигнале равна $\log_2 8$ = 3 битам независимо от количества нулей и единиц в управляющем сигнале. Если управляющий сигнал состоит, например, из пяти нулей и единиц, то информацию несут только три, остальные символы являются избыточными, но могут быть использованы для страховки, если, например, при передаче произойдут искажения.

Шахматная игра является, конечно, простейшим примером, ибо в ней мы знаем до конца все разнообразие возможных ходов каждой фигуры. Так бывает далеко не всегда. Например, наследственная информация, заключенная в хромосомах зародышевых клеток, является, безусловно управляющей информацией. Она определяет путь, по которому пойдет развитие зародыша и образование взрослого организма, но оценка количества наследственной информации затрудняется тем, что мы не знаем "количества разнообразия" возможных путей развития зародыша.

Разнообразие различных форм и видов информации - как уже известных, так и тех, которые будут введены и исследованы в дальнейшем - заставляют с особым вниманием отнестись к тому, что объединяет между собой эти различные формы и виды. Этим объединяющим является то, что, согласно определению, любая форма информации является мерой сигнального взаимодействия между материальными системами. Легко убедиться, что определение понятия "информация", введенное в § 6, охватывает все известные формы и виды информации, создавая тем самым единую основу для их анализа.

§ 9. "ДЕМОН МАКСВЕЛЛА"

Под названием "Демон Максвелла" известно "гипотетическое существо", участвующее в мысленном эксперименте, описанном Д.К.Максвеллом в 1871 г. Этот мысленный эксперимент интересен тем, что он, казалось бы, опровергает универсальность второго начала термодинамики.

Действительно, если два изолированных от внешнего мира объема газа сообщаются между собой хотя бы через маленькое отверстие (рис.12), то согласно второму началу термодинамики температуры, а тем самым и средние скорости молекул в обоих объемах газа должны со временем выравниваться. Максвелл предложил следующий мысленный эксперимент: снабдим отверстие дверцей, около которой посадим гипотетическое существо (можно условно назвать его "демоном"), которое способно видеть отдельные молекулы и различать их скорости. Пусть теперь этот "демон" открывает дверцу только тогда, когда быстрая ("горячая") молекула летит из левого сосуда в правый, или когда медленная ("холодная") молекула летит из правого сосуда в левый, а во всех остальных случаях держит дверцу закрытой. В результате работы "демона" в правом сосуде соберутся более быстрые ("горячие") молекулы, а в левом - более медленные ("холодные"), и энтропия газа уменьшается. Хотя эксперимент Максвелла носит мысленный характер, мы должны уметь объяснить его, поскольку существование хотя бы и мысленного эксперимента, противоречащего физическим законам, говорит о недостаточно четком понимании этих законов и нуждается в объяснении.

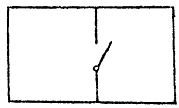

Рис.12

Попробуем объяснить парадокс "демона Максвелла", исходя из понимания информации как о мере сигнального взаимодействия между материальными системами. Для того чтобы сортировать "горячие" и "холодные" молекулы "демон" должен получить информацию о величине и направлении скорости молекулы, подлетающей к дверце, которой он управляет. Согласно определению, информация - это мера взаимодействия между материальными системами, а всякое реальное взаимодействие (в том числе и сигнальное, информационное взаимодействие) - это необратимый процесс, сопровождающийся увеличением энтропии. Рассмотрим количественную сторону. Если температура газа в сосудах (а с ним и температура "демона") равна T, то для того, чтобы быть замеченным на фоне тепловых флюктуаций, сигнал о каждой отдельной молекуле должен обладать энергией, большей, чем $\frac{1}{2}kT$ (энергия на степень свободы в равновесном распределении) в то время как приращение свободной энергии в результате "отсортировки" одной молекулы составляет лишь доли от $\frac{1}{2}kT$; таким образом, увеличение энтропии в результате наблюдения за молекулами превышает уменьшение энтропии, достигаемое вследствие их сортировки. Кажущееся противоречие парадокса Максвелла второму началу термодинамики связано с недоучетом того обстоятельства, что получение информации - это не идеальный, а материальный процесс, а все реальные процессы - необратимы и сопровождаются увеличением энтропии. Как правильно подчеркивает Бриллюен, "мы ничего не можем получить даром, даже наблюдения".

Таким образом, парадокс "демона Максвелла" получает простое объяснение и не требует, как это делается в [2], отождествления информации с отрицательной энтропией, "негэнтропией".

Характеризуя своего "демона", осуществляющего сортировку молекул, Д.К.Максвелл определяет его как "существо со столь изощренными способностями, что он может следить за движением каждой молекулы и делать то, к чему мы в настоящее время (Д.Максвелл писал это в 1871 г.) неспособны". Таким образом, "демон" определялся как существо, преодолевающее ограниченность человеческих возможностей того времени. Однако ограничения на человеческие возможности носят двоякий характер. Одни из них временны и определяются лишь несовершенством человеческих знаний или техники. Другие определяются законами природы и не могут быть преодолены техническими ухищрениями. Во времена Максвелла не существовало приборов, способных измерить скорость молекул. Сейчас такие приборы есть. Однако любое измерение по-прежнему остается материальным и необратимым процессом, сопровождающимся увеличением энтропии. Это увеличение энтропии будет препятствовать функционированию "демона Максвелла". Между информацией и энтропией, таким образом, действительно существует связь, однако совсем не такая тесная и не столь прямолинейная, как это утверждал "негэнтропийный принцип информации".

Эта связь определяется только тем, что любой сигнал, несущий информацию, должен обладать определенной энергией, чтобы быть принятым, причем величина этой энергии зависит от конкретных свойств приемника и может быть самой различной. В результате приема сигнала - необратимого процесса - происходит увеличение энтропии, но степень этого увеличения зависит от конкретных свойств приемника и может быть самой различной.

Интересна роль "демона Максвелла" в истории физики. К тому времени, когда Максвелл впервые написал о "демоне" (1871 г.) второе начало термодинамики было уже общепризнанным физическим законом, было уже ясно, что энтропия изолированной системы не может убывать и "вечный двигатель второго рода" - т.е. машина, уменьшающая энтропию, также невозможен, как и "вечный двигатель первого рода" - т.е. машина, создающая энергию. Между тем Максвелл описал мысленный эксперимент с гипотетическим существом, которое он назвал "демоном" и этот эксперимент опровергал второе начало, поскольку "демон", сортируя молекулы, уменьшал энтропию системы.

Конечно, от второго начала термодинамики "демон Максвелла" не заставил отказаться, но мысленный эксперимент Максвелла указывал на то, что в физике его времени существовал пробел и этот пробел заключался в том, что еще не была до конца осознана материальность процесса получения информации о скорости молекул, в связи с чем любое измерение скорости (в том числе и "демоном"), любое получение информации сопровождается увеличением энтропии, что и не позволяет "демону" нарушить второе начало. Не опровергнув второго начала термодинамики, "демон Максвелла" долгие годы тревожил воображение физиков и способствовал выработке правильного понимания процессов, происходящих при измерении.

§ 10. ЗАВИСИМОСТЬ ТОЧНОСТИ СИСТЕМ УПРАВЛЕНИЯ ОТ ИНФОРМАЦИИ О ВОЗМУЩАЮЩИХ ВОЗДЕЙСТВИЯХ И ВЫХОДЕ СИСТЕМЫ

В настоящем разделе мы рассмотрим линейные системы управления и проанализируем - каким образом зависит точность управления от располагаемой информации о возмущающих воздействиях.

Рассмотрим простейшую систему управления, математической моделью, которой является уравнение:

$$\mathcal{D}x = u + \varphi(t), \tag{38}$$

где $\mathcal{D} = \frac{d}{dt}$; x - регулируемая величина, u - управляющее воздействие, $\varphi(t)$ – возмущающие воздействие, случайный процесс, распределенный по нормальному закону и имеющий среднее значение равное нулю, дисперсию $\langle \varphi^2 \rangle = k^2$ и спектральную плотность мощности $S_\varphi = \frac{2}{\pi}\frac{k^2}{1+\omega^2}$. Управление формируется в функции от величины $x(t)$, – выхода системы управления –, но известной не точно, а с погрешностью $\psi(t)$ – т.е., фактически, $u = u(x+\psi)$, где $\psi(t)$ – случайный процесс, распределенный по нормальному закону, со средним значением, равным нулю, дисперсией $\langle\psi^2\rangle = 1$ и спектральной плотностью мощности $S_\psi = \frac{2}{\pi}\cdot\frac{1}{1+\omega^2}$ (таким образом, информация о переменной $x(t)$ является не точной). Как показано в работе [I], в этом случае оптимальный закон регулирования $u_{опт} = u(x+\psi)$, обеспечивающий наивысшую возможную точность регулирования, и само предельно возможное значение этой точности, зависят только от отношения дисперсии возмущающего воздействия к дисперсии погрешностей измерения, а от других характеристик информации о процессах $\varphi(t)$ и $\psi(t)$ – не зависят (в частности – не зависят и от энтропии распределения вероятностей, определяющей количество статистической информации по Шеннону). Действительно, как показано в [I], имеем:

$$u_{опт} = -[kD + k(k+1)](x+\psi), \tag{39}$$

$$\langle x^2 \rangle_{min} = 1 - \frac{1}{(k+1)^2}. \tag{40}$$

Рассмотрим теперь ту же систему (38) для случая, когда погрешностью измерения $x(t)$ можно пренебречь, но необходимо учитывать ограничения на управление и поэтому критерием качества является $J = m^2\langle x^2\rangle + \langle u^2\rangle$, где m^2 – множитель Лагранжа, определяемый из условия $\langle u^2\rangle \le N_u$. Если возмущающие воздействия $\varphi(t)$ распределены по нормальному закону, то закон регулирования, обеспечивающий устойчивость замкнутой системы и минимум критерия качества, а также и само значение минимума, зависят только от спектральной плотности мощности процесса $\varphi(t)$, а от другой информации о процессе $\varphi(t)$ – не зависят; как показано в работе [I], если $S_\varphi = \frac{2}{\pi}\cdot\frac{\alpha}{\alpha^2+\omega^2}$, то

$$u_{опт} = -\left(\frac{m}{\alpha} \mathcal{D} + m \frac{\alpha + m}{\alpha} \right) x \, , \tag{41}$$

$$J_{min} = \frac{m(m+2\alpha)}{(m+\alpha)^2} \, . \tag{42}$$

Если мы заранее, еще при $t = 0$ располагаем информацией о всех будущих значениях возмущающего воздействия, то можно реализовать программное управление, и в этом случае качество системы управления может быть улучшено. Предельное значение критерия качества (абсолютный минимум) равно

$$J_{абс\ min} = \frac{m}{\alpha + m} \, , \tag{43}$$

но достижимо оно лишь в том случае, если $\varphi(t)$ заранее известно нам. Для всех $\alpha > 0$ будет иметь место неравенство $J_{абс\ min} < J_{min}$, которое при $\alpha = 0$ переходит в равенство: $J_{абс.\ min} = J_{min}$. Получается, что при $\alpha = 0$ информация о корреляционной функции процесса $\varphi(t)$ эквивалентна полному знанию процесса. Однако рассмотрев процесс, корреляционная функция которого является экспонентой: $K_\varphi = e^{-\alpha\tau}$ убеждаемся, что при $\alpha = 0$ будет $\varphi(t) = 1$ или $\varphi(t) = -1$ – т.е. при $\alpha = 0$ вся информация о процессе $\varphi(t)$ действительно заключена (с точностью до знака) в его корреляционной функции. При $\alpha > 0$ корреляционная функция полной информации о процессе $\varphi(t)$ не содержит.

Мы убеждаемся, что законы управления и достижимая точность действительно зависят от количества информации о выходе системы $x(t)$, возмущающем воздействии $\varphi(t)$ и погрешности измерения $\psi(t)$, но это – совсем не та информация с которой имеет дело техника связи и количество которой измеряется по формуле Шеннона (37). Вряд ли нам удалось достичь чего-либо хорошего, если бы при анализе систем управления мы довольствовались бы определением количества информации по Шеннону.

Понимание того, что существуют другие виды, другие формы информации, отличные от статистической, позволяет более обоснованно подойти к решению задач управления. Вместе с тем – и это надо прямо сказать – еще далеко не ясно, какие именно характеристики информации о возмущающих воздействиях следует стремиться получить для улучшения качества управления. Если система линейна, а критерий – среднеквадратичный, то исключительную роль играет корреляционная функция возмущающего воздействия. Достижимая точность управления зависит только от нее. Однако возможны и другие варианты.

Рассмотрим снова систему (38) с критерием качества $J = \max|x(t)|$ при ограничении $|u| \le 1$. Пусть для $0 \le t \le t_1$ будет $|\varphi(t)| \le 1$, при $t_1 \le t \le t_2$ будет $\varphi(t) > 1$, а при $t > t_2$ снова будет $|\varphi(t)| \le 1$. Тогда для обеспечения минимума критерия качества нужно заранее, при $t_0 < t_1$ переключить управление от $u = -\varphi$ к значению $u = -1$, причем момент переключения t_0 нужно выбрать из условия:

$$2\int_{t_0}^{t_1}(1-\varphi)\,dt = \int_{t_1}^{t_2}(\varphi-1)\,dt \tag{44}$$

– т.е. для выбора оптимального управления нужно в момент $t = t_0$ располагать сведениями о величине интегралов (44). Еще раз убеждаемся, что для правильного формирования управления нужна информация, но не статистическая, не шенноновская. Более подробные данные об информации, требующейся для различных задач управления – смотри в работе [I].

Разумеется, теория, изложенная в предыдущих разделах, не позволяет сама по себе дать рецепт для выбора нужной для целей управления различными объектами информации и для оценки ее количества, но она дает хотя бы общую ориентировку в проблеме. Если считать, что никакой другой информации, кроме измеряемой по формулам Шеннона или Хартли не существует (а ведь долгие годы так и считали), то вообще нельзя понять, почему в области управления нам нужны знания и информация совсем другого характера, чем в области теории связи. Понимание этого обстоятельства – первый шаг к построению полной количественной теории.

Отметим, что материал, изложенный в настоящей книге, был собран давно и неоднократно обсуждался (в том числе на Всесоюзных конференциях), начиная с 1965 г. Настоящая книга написана в 1968г. и несколько дополнена в 1987 г. В издательство рукопись была передана в 1974 г., но печатание задержалось по причинам, не зависящим от автора. Об этом необходимо упомянуть потому, что в 1981г. в Чехословакии вышла книга Я.Черни под очень похожим названием, хотя и с другим содержанием. Поскольку рукопись настоящей книги начала свой путь к читателю еще в 1974 г., автор счел возможным сохранить ее первоначальное название.

Заключение

Анализ понятий упорядоченности, энтропии, информации, проведенный нами в ходе предыдущего изложения, позволяет сделать следующие выводы:

I. Термодинамическая энтропия не является мерой неупорядоченности системы. Она является характеристикой вероятности ее состояния и позволяет предсказывать направленность процессов, происходящих в изолированных системах.

Предоставленная самой себе изолированная система переходит от менее вероятных состояний (состояний с меньшей энтропией) к более вероятным состояниям, однако второе начало термодинамики не означает, что в изолированных системах обязательно существует тенденция к возрастанию неупорядоченности, к переходу системы в наиболее беспорядочное, хаотическое состояние.

В зависимости от конкретных условий, в зависимости от характера сил, действующих на частицы и компоненты системы, протекающие в ней процессы могут приводить как к увеличению неупорядоченности, так и к росту порядка.

Широко распространенное популярное представление об энтропии как о мере неупорядоченности помогает на первой стадии изучения энтропии, позволяя составить о ней наглядное представление. Оно полезно во многих случаях, поскольку существует весьма большое количество систем, в которых рост энтропии сопровождается и ростом неупорядоченности, однако распространение представления об энтропии как мере неупорядоченности "на все случаи жизни" может привести к ошибкам и породить ошибочные представления о направленности процессов в окружающем нас мире. Более правильно понимать второе начало термодинамики как закон перехода от менее вероятных состояний к более вероятным, который - в зависимости от конкретных условий - может сопровождаться как ростом, так и уменьшением неупорядоченности.

2. Энтропия распределения вероятностей, вычисляемая по формуле

$$H = -\sum p_i \log p_i$$

является математической характеристикой распределения вероятностей исходов опыта, характеризующей трудность предсказания его исхода. Сходство между этой формулой и формулой для вычисления энтропии в физике недостаточно для отождествления энтропии распределения вероятностей, используемой в теории информации, и термодинамической энтропии. Смешение этих понятий может затруднить понимание и привести к ошибкам.

3. Под единым понятием "информация" скрывается несколько различных форм информации, имеющих между собой не так уж много общего.

Помимо информации, изучаемой в классической теории связи, мерой количества которой служит энтропия распределения вероятностей принимаемых символов и которую можно называть статистической информацией существуют и имеют не меньшее значение другие виды, другие формы информации. Одним из видов является целевая информация, когда принимаемые сигналы используются для вполне определенной цели и количество информации измеряется по изменению целевой функции в результате приема сообщения. Другим видом информации является смысловая информация, не предназначенная для определенной цели, а расширяющая объем знаний, изменяющая "тезаурус" получателя. Возможно выделение и других видов информации, отличных от перечисленных. Общим для всех видов информации является то, что они являются характеристиками, мерами различных форм сигнальных взаимодействий между материальными системами.

Важно отметить, что наиболее интересные, наиболее увлекательные приложения теории информации связаны с ее целевой и смысловой формами. В то же время достаточно полные и совершенные методы вычисления количества информации разработаны пока только для статистической информации. Поэтому приложения теории информации в области психологии, теории искусства и литературы пока еще не очень успешны. В дальнейшем, по мере разработки методов вычисления количества различных видов информации следует ожидать новых интересных открытий.

Литература

1. Абдуллаев Н.Д., Петров Ю.П. Теория и методы проектирования оптимальных регуляторов. Л., 1986.
2. Бриллюэн Л. Наука и теория информации. М., 1960.
3. Винер Н. Кибернетика. М., 1958.
4. Гленсдорф П., Пригожин И. Термодинамическая теория структуры, устойчивости и флюктуаций. М., 1974.
5. Гроот С., Мазур П. Неравновесная термодинамика. М., 1964.
6. Котоусов Л.С. Принцип минимума производства энтропии и его применение в практических задачах// Проблемы современной физики. Л., 1974.
7. Лернер В.С. Применение подхода теории неравновесных процессов некоторым задачам управления. Кишинев, 1968.
8. Петров Ю.П. К вопросу о принципе минимума скорости возрастания энтропии в стационарном режиме// Биофизика. 1966. Т.11, № 5. С.926-928.
9. Петров Ю.П. Вариационные методы теории оптимального управления. 2-е изд. Л., 1977.
10. Пригожин И.Р. Время, структура и флюктуации// Успехи физических наук. 1980. Т.131. Вып.2. С.185-207.
11. Стеблев Ю.И. Условия минимальных потерь в теории поверхностного эффекта// Электричество. 1983, № 6. С.62-65.
12. Теория передачи электрических сигналов при наличии помех/ Под ред.Н.А.Железнова, М., 1953.
13. Шеннон К. Работы по теории информации и кибернетике. М., 1963.

Оглавление

Петров Юрий Петрович

ИНФОРМАЦИЯ И ЭНТРОПИЯ В КИБЕРНЕТИКЕ

Учебное пособие

Доп. темплан 1989 г.

Подписано в печать с оригинала-макета 25.10.89. Ф-т 60х90/16.
Бум. тип. № 3. Печать офсетная. Усл. печ. л. 5,25. Усл. кр.-отт. 5,25.
Уч.-изд. л. 5,0. Тираж 500 экз. Заказ № 492. Цена 15 коп.

Редакционно-издательский отдел Ленинградского университета.
199034, Ленинград, Университетская наб., 7/9.

Печатно-множительная лаборатория Ленинградского университета.
199034, Ленинград, наб. Макарова, 6.