

INFORMATION,
THERMODYNAMICS,
AND THE STRUCTURAL
PRINCIPLES
OF BIOLOGICAL
SYSTEMS

L. A. BLUMENFELD

The article is devoted to the analysis of interrelations between information and thermodynamics. The quantitative correlation between entropy and quantitative information, which is necessary for the full description of a system is discussed. Also examined are the questions of meaning of biological regulation and the concept of "construction" of the living systems and their components.

Статья посвящена анализу взаимосвязи между информацией и термодинамикой. Рассмотрены количественные соотношения между энтропией и количеством информации, необходимой для полного описания системы. Обсуждаются вопросы о смысле биологической упорядоченности и понятии "конструкции" живых систем и их компонентов.

© Блюменфельд Л.А., 1996

ИНФОРМАЦИЯ, ТЕРМОДИНАМИКА И КОНСТРУКЦИЯ БИОЛОГИЧЕСКИХ СИСТЕМ

Л. А. БЛЮМЕНФЕЛЬД

Московский государственный университет
им. М.В. Ломоносова

ЧТО ТАКОЕ ИНФОРМАЦИЯ?

“Я бросал игральный кубик пять раз и каждый раз выпадало два очка”. “Из колоды, содержащей 32 карты, я наугад выбрал одну, и она оказалась дамой пик”. “Я взвесил образец на весах с точностью ± 10 грамм. Вес оказался равным 100 грамм. Я повторил взвешивание на весах с точностью ± 1 грамм, и вес оказался равным 99 грамм”. В каком из этих трех случаев я получил больше информации? Во всех этих примерах речь идет о получении сообщения, ограничивающего возможность выбора вариантов значения некоторой величины. В этом смысле любое измерение ничем не отличается от получения сообщения по любым каналам связи.

Проблема оценки количества информации, содержащегося в сообщении, была решена в 1949 году [1].

Для упрощенной ситуации: когда может произойти $1/p$ событий с одинаковыми вероятностями $p \leq 1$ для каждого из них, количество информации I определяется формулой

$$I = \log_2 \frac{p_1}{p}, \quad (1)$$

где p – априорная вероятность некоторого события (то есть вероятность до получения сообщения), а p_1 – вероятность после получения сообщения. Если сообщения достоверны и однозначны, то $p_1 = 1$

$$I = -\log_2 p. \quad (2)$$

В качестве единицы информации I принимают количество информации в достоверном сообщении о событии, априорная вероятность которого равна $1/2$. Эта единица получила название “бит” (от английского binary digits).

В первом из приведенных выше примеров $p = (1/6)^5 = 1/7776$ и $I = \log_2 7776 \approx 12,9$ бит; во втором примере $I = -\log_2(1/32) = 5$ бит. В третьем примере повышение точности измерения в 10 раз приводит к десятикратному сужению интервала возможных значений измеряемой величины и, следовательно, к уменьшению в 10 раз вероятности получения измеренного значения: $I = -\log_2(1/10) \approx 3,33$ бит.

ИНФОРМАЦИЯ И ЭНТРОПИЯ

За 20 лет до появления работ Шеннона анализом информации занимался венгерский физик Сциллард [2] в связи с решением одного термодинамического парадокса, предложенного Максвеллом еще в прошлом веке [3]. Смысл парадокса Максвелла заключается в следующем (рис. 1). Изолированная система, состоящая из разделенного на две части резервуара с газом и с дверцей в перегородке, содержит также “демона” (существо или автомат), наделенного способностью отличать быстрые молекулы от медленных. Демон открывает дверцу только в том случае, если к ней справа подлетает быстрая молекула. Поэтому газ в левой части резервуара будет нагреваться, а в правой — охлаждаться. Таким образом, в изолированной системе тепло будет переходить от холодного тела к горячему с понижением энтропии системы в противоречии со вторым законом термодинамики.

Сциллард [2], рассмотрев один из упрощенных вариантов парадокса Максвелла, обратил внимание на необходимость получения информации о молекулах и открыл связь между информацией и термодинамическими характеристиками. В дальнейшем решение парадокса Максвелла было предложено многими авторами (см., например, [4]). Смысл всех решений заключается в следующем: информацию нельзя получать бесплатно. За нее приходится платить энергией, в результате чего энтропия системы повышается на величину, по крайней мере равную ее понижению за счет полученной информации.

Пусть задано макроскопическое состояние некоторой системы, то есть с определенной степенью точности указаны значения таких параметров, как объем, давление, температура, химический состав и т.п. Каждому макросостоянию системы соответствует набор микросостояний. В микросостоянии прецизированы (точно заданы) состояния всех частиц, входящих в систему. Для любой макросистемы при температуре выше абсолютного нуля число микросостояний W , соответствующих данному макросостоянию, огромно. W называется статистическим весом или термодинамической вероятностью данного макросостояния. Согласно основному постулату

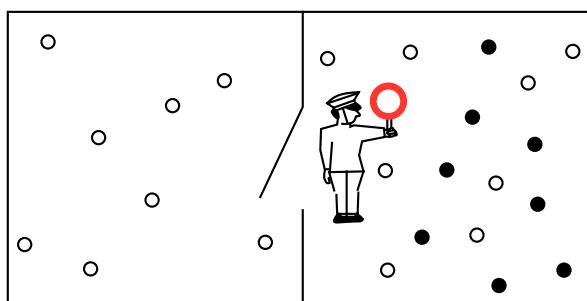


Рис. 1. Демон Максвелла.

статистической физики, все W микросостояний, соответствующие одному макросостоянию, имеют одинаковую априорную вероятность. *Знать микросостояние системы — значит знать о системе все!*

Величина W непосредственно связана с энтропией. По формуле Планка–Больцмана

$$S = k \ln W, \quad (3)$$

где размерная постоянная Больцмана $k = 1,38 \times 10^{-16}$ эрг/град или $3,31 \cdot 10^{-24}$ эе (эе — энтропийная единица, $1 \text{ эе} = 1 \text{ кал/град}$). Рассчитаем, какое количество информации надо получить о системе, находящейся в данном макросостоянии, чтобы однозначно определить ее микросостояние. Иначе говоря, какого количества информации недостает для полного описания системы в заданном макросостоянии?

Пусть микросостояние определено путем измерений или расчетов (на самом деле сделать это нельзя). До определения вероятность того, что макроскопическая система находилась именно в этом микросостоянии, была равна $1/W$, а после определения стала равной единице. Полученное количество информации

$$I = -\log_2 \frac{1}{W} = \log_2 W. \quad (4)$$

Формулы (3) и (4) совпадают с точностью до постоянного размерного множителя. Величины I и S существенно идентичны. Ситуация здесь та же, что и для соотношения между массой и энергией: $E = mc^2$, где роль размерного множителя играет c^2 . Аналогична ситуация и в случае соотношения между частотой и энергией кванта света, где размерным множителем является постоянная Планка: $\epsilon = \hbar \omega$. Энтропия системы в данном макросостоянии есть количество информации, недостающее до ее полного описания. Чтобы перейти от количества информации в битах к энтропии в энтропийных единицах, необходимо перейти от логарифма при основании 2 к натуральному логарифму и умножить на k :

$$S(\text{эе}) = 2,3 \cdot 10^{-24} I \text{ бит.} \quad (5)$$

СКОЛЬКО СТОИТ БИОЛОГИЧЕСКАЯ УПОРЯДОЧЕННОСТЬ?

Почти во всякой книге, посвященной теоретической биологии и биофизике, можно встретить слова об удивительной упорядоченности биологических структур на всех уровнях от макромолекул до человека. Рассуждения об антиэнтропийных тенденциях живого, об антиэнтропийном характере биологической эволюции стали общим местом не только у биологов и философов, но и у физиков. Эти утверждения основаны на чувстве удивления, которое испытывают ученые, наблюдающие биологические структуры и процессы. Для утверждения, что упорядоченность одной системы выше, чем другой, нужно эти упорядоченности измерить.

Измеримая, выражаемая числом упорядоченность может относиться лишь к макросостоянию системы. В распоряжении физики имеется пока лишь одна величина, характеризующая степень упорядоченности — энтропия. Оценим поэтому изменения энтропии, связанные с возникновением биологической организации.

Упорядоченность построения многоклеточного организма из клеток

Тело человека содержит примерно 10^{13} клеток. Допустим, что среди них нет ни одной пары одинаковых и что ни одну пару нельзя поменять местами без нарушения функционирования организма. Это значит, что относительное расположение клеток в теле человека однозначно. Количество информации, необходимой для построения такой единственной структуры из $10^{13}!$ возможных,

$$I = \log_2(10^{13}!) \approx 10^{13} \log_2 10^{13} \approx 4 \cdot 10^{14} \text{ бит.} \quad (6)$$

Отсюда следует, что понижение энтропии при построении организма человека из клеток составит

$$\Delta S \approx 2 \cdot 10^{-24} \cdot 4 \cdot 10^{14} \approx 10^{-9} \text{ эе.} \quad (7)$$

При испарении одного грамма воды энтропия повышается примерно на 1 эе. Таким образом, понижение энтропии при переходе от хаотически расположенных клеток к организму человека численно равно повышению энтропии при испарении 10^{-9} грамм воды. Принятые постулаты об отсутствии одинаковых клеток и о невозможности их перестановок лишь увеличили рассчитанное по формуле (6) количество требуемой информации, так что ΔS на самом деле гораздо меньше.

Упорядоченность построения клетки из биополимеров

Количество молекул биополимеров (белков, нуклеиновых кислот, полиуглеводов и др.) в одной клетке составляет в среднем 10^8 . Допустим снова, что все молекулы различны, а их относительное расположение уникально. Количество информации, необходимой для построения одной клетки из готовых биополимеров,

$$I \approx 10^8 \log_2 10^8 \approx 2,6 \cdot 10^9 \text{ бит,} \quad (8)$$

а для всех клеток в организме человека $2,6 \cdot 10^{22}$ бит, что соответствует понижению энтропии примерно на $6 \cdot 10^{-2}$ эе.

Упорядоченность построения белков и ДНК из мономеров

Организм взрослого человека содержит около 7 кг белков и 150 г ДНК, что соответствует $\approx 3 \cdot 10^{25}$ аминокислотных и $\approx 3 \cdot 10^{23}$ нуклеотидных остатков. Для создания единственной последовательности из $20^3 \cdot 10^{25}$ возможных, для белка необходимо $\approx 10^{26}$ бит. Для ДНК необходимо $\approx 6 \cdot 10^{23}$ бит. В пересчете на

энтропию получаем для белков и ДНК 300 и 1,4 эе соответственно.

Таким образом, упорядоченность биологической организации человеческого тела “стоит” 301,5 эе и подавляющий вклад вносит упорядоченное распределение аминокислотных остатков в белках. Понижение энтропии при возникновении такой биологической организации с легкостью компенсируется тривиальными физическими и химическими процессами. Повышение энтропии на 300 эе обеспечивается испарением 170 г воды.

Эти оценки показывают, что возникновение и усложнение биологической организации происходит практически “бесплатно”. Все разговоры об антиэнтропийных тенденциях биологической эволюции основаны на недоразумении. Согласно физическим критериям, любая биологическая система упорядочена не больше, чем кусок горной породы того же веса.

Хотя эти оценки не могут быть оспорены, они вызывают чувство неудовлетворенности. Ведь без всяких расчетов поразительная упорядоченность биологических структур и процессов совершенно очевидны. Если физика говорит, что эта упорядоченность ничего не стоит, то хочется ответить: тем хуже для физики. Интуитивно чувствуется, что биоструктуры обладают особой упорядоченностью, измерять которую в энтропийных единицах, конечно, можно, но многое для понимания особенностей живой материи при этом не стоит ожидать.

В основе ощущения особой упорядоченности биологических структур лежит то обстоятельство, что она имеет смысл.

Осмысленно то, что имеет цель.

СМЫСЛ БИОЛОГИЧЕСКОЙ УПОРЯДОЧЕННОСТИ

Исследуя природу, естествоиспытатель задает ей вопросы. При изучении мертвой материи возможны только два принципиальных вопроса: как и почему? Ответы на первый вопрос дают описания строения исследуемых объектов, последовательности событий, с ними происходящих (например, в ходе химических превращений, при прохождении света через различные среды, во время землетрясений и т.п.). Ответы на второй вопрос дают физические и химические законы. Почему кристалл поваренной соли имеет кубическую структуру с данным расположением ионов Na^+ и Cl^- друг относительно друга? Ответ: потому что расчеты с использованием законов физики показывают, что такая структура соответствует минимуму свободной энергии системы, то есть наиболее стабильна. Кстати, требование минимума свободной энергии значит, что структура кристалла не может быть идеально правильной. Вхождение в уравнение для свободной энергии $\Delta F = \Delta U - T\Delta S$ члена $T\Delta S$ приводит к появлению равновесной концентрации дефектов кристаллической решетки.

При исследовании объектов живой природы и объектов, сделанных живой материей (главным образом в результате деятельности человека), можно задать третий вопрос: для чего? Собственно говоря, возможность задать такой вопрос и свидетельствует, что живая материя, ее компоненты и объекты, ею изготовленные, имеют смысл. Нельзя спросить: для чего кристалл NaCl имеет кубическую симметрию? Однако можно спросить: для чего молекула гамма-глобулина построена так, а не иначе? И получить ответ: для того чтобы осуществлять функции иммунной защиты и предотвращать гибель организма, уменьшая вероятность гибели вида.

Попробуем разобраться в том, что такое "смысл" биологической упорядоченности и какими свойствами должны обладать системы, упорядоченность которых имеет смысл. Впервые четко поставил эти вопросы и попытался на них ответить Кастлер [5]. Приведу некоторые примеры из этой книги.

Пусть имеется сейф с замком. Есть набор цифр от 0 до 9 и устройство случайного выбора, позволяющее выбрать тройку цифр. Выберем три цифры и введем их в качестве кода в замок сейфа. До этого любые мыслимые последовательности из трех цифр ничем друг от друга не отличались. Все они не имели смысла. Введение случайно отобранный тройки цифр в замок сейфа сделало ее осмысленной. Смысл ее заключается в том, что она открывает сейф, а другие нет.

Другой пример возникновения осмысленной упорядоченности. Пусть в обширном водном резервуаре растворено большое количество нуклеотидов – мономеров, из которых построена ДНК. Присутствуют четыре сорта нуклеотидов: на основе аденина (А), гуанина (Г), цитозина (Ц) и тимина (Т). Между нуклеотидами могут идти реакции конденсации с образованием однонитевых ди-, три- и т.д. полинуклеотидов. Например,



Скорость распада ди-, три-, ... полинуклеотидов значительно больше скорости их образования, и равновесие в этих реакциях будет сдвинуто в сторону мономеров. Концентрации полинуклеотидов будут невелики и тем меньше, чем длиннее цепочка. Небольшое число полимерных нитей разной длины, которые всегда будут присутствовать в растворе в динамическом равновесии с огромным избытком мономеров, должно иметь совершенно случайные последовательности нуклеотидов (скорости реакций распада и синтеза для всех нуклеотидов практически одинаковы), и априорные вероятности всех последовательностей будут одинаковы. Ясно, что та или иная реализующаяся последовательность, та или иная возникающая упорядоченность не будут иметь смысла.

Однако благодаря особым химическим свойствам нуклеотидов (возможность образования водородных связей между ними), помимо перечисленных выше реакций образования однонитевых цепочек, могут идти реакции присоединения цепочки других нуклеотидов и образование новой цепочки, связанной с первой. Этот процесс называется матричным синтезом. Последовательность нуклеотидов в новой цепочке полностью определяется последовательностью в исходной цепочке: против А всегда стоит Т, а против Г – Ц:

А Т Г Ц Т А Ц Г Г А...

Т А А Ц Г А Т Г Г Ц Т...

Матричный синтез

Возникающая в результате матричного синтеза молекула двунитевого полимера гораздо стабильнее молекулы однонитевого. По достижении достаточной длины она образует двойную спираль, практически не распадается и может успеть "редублицироваться", нарастив на обе нити соответствующие нуклеотиды:

Т А А Ц Г А Т Г Г Ц Т...

А Т Г Г Ц Т А Ц Г Г А... $\xrightarrow{2} \begin{pmatrix} \text{Т А А Ц Г А Т Г Г Ц Т...} \\ \text{А Т Г Г Ц Т А Ц Г Г А...} \end{pmatrix}$

Редубликация

Первая двунитевая молекула образуется в результате случайного и весьма маловероятного процесса: на одной из однонитевых полимерных молекул успевает до ее распада пройти матричный синтез. Последовательность нуклеотидов в этой однонитевой молекуле могла быть любой. Однако после того, как двунитевая структура образовалась, ситуация резко изменилась. Последовательность, реализованная в таком долгоживущем двунитевом полимере, приобрела смысл. Этот смысл состоит в том, что эта последовательность в стабильной и способной к редубликации молекуле существует, а другие возможные последовательности – нет. В системе будет быстро возрастать концентрация полимеров именно с такой, теперь уже особой, последовательностью. Случайные отклонения ("ошибки") от "правильной" последовательности будут также воспроизводиться и дадут начало самостоятельным системам, конкурирующим с исходной за наличные запасы мономеров. Таким образом, благодаря запоминанию случайного выбора возникла упорядоченность, имеющая смысл, возникла система, способная создавать осмысленную информацию.

Этот пример стабильной самовоспроизводящейся системы нуклеиновых кислот ни в коем случае не претендует на то, чтобы его связывали с происхождением жизни на Земле. Это просто попытка продемонстрировать наиболее существенные

характеристики процесса создания осмысленной упорядоченности на примере биополимеров с использованием некоторых хорошо известных их свойств.

В обоих приведенных примерах речь шла именно о создании новой информации. До того, как полимер со случайной последовательностью нуклеотидов образовал стабильную двунитевую структуру, или до того, как случайная последовательность трех цифр была введена в замок сейфа, информации о том, что эти последовательности “лучше” других, просто не существовало. Информация была создана, сотворена. Непредсказуемое превратилось в неизбежное (Пьер Буле, цитировано по [5]).

Системы, создающие осмысленную упорядоченность, обладают одним общим свойством: они содержат компоненты, конструкции, продолжительность жизни которых превышает время одного цикла работы системы. Для системы нуклеиновых кислот это значит, что двунитевой полимер не распадается до редубликации, а для другого примера – что замок сейфа не разрушится до того, как будет испытана хотя бы одна тройка цифр. Требование наличия долгоживущих, медленно релаксирующих конструкций обязательно для живой материи. Нельзя построить живое на основе газовой фазы!

Таким образом, понятие конструкции становится весьма важным при анализе функционирования

живых систем и их компонентов. С позиции статистической физики наличие конструкции означает наличие границ между областями фазового пространства, пересечение которых запрещено для figurативных точек статистической системы в течение данного промежутка времени. Эта проблема выходит за рамки настоящей статьи.

ЛИТЕРАТУРА

1. *Shannon C.E., Weaver W. The Mathematical Theory of Communication.* Univ. of Illinois Press, 1949.
2. *Szillard L. // Z. Physik.* 1929. V. 53. P. 840.
3. *Maxwell J.C. Theory of Heat.* London, 1971.
4. *Бриллюен Л. Наука и теория информации.* М.: Физматгиз, 1960.
5. *Кастлер Г. Возникновение биологической организации.* М.: Мир, 1967.

* * *

Лев Александрович Блюменфельд, доктор химических наук, действительный член Российской академии естественных наук, профессор кафедры биофизики Физического факультета Московского государственного университета им. М.В. Ломоносова, главный научный сотрудник Института биохимической физики. Автор около 300 научных работ и 5 книг.