

О понятии информационного взаимодействия, 1: допсихический уровень

Н.А.Кузнецов, В.А.Любецкий, А.В.Чернавский

*Институт проблем передачи информации РАН,
101447, Москва, Большой Каретный переулок, 19, Россия,
e-mail: lyubetsk@iitp.ru*

Поступила в редколлегию 29.01.2003

Аннотация—В нескольких статьях, начиная с этой, авторы предполагают изложить соображения, относящиеся к понятиям информационного взаимодействия и информационной, сигнальной, языковой, семантической составляющим в живом. Сейчас, когда такого рода представления только зарождаются, уже просто отбор и систематизация материала на эту тему представляют, на наш взгляд, некоторый интерес. По ходу изложения упоминаются отдельные модели, но их окончательные формулировки и, тем более, математическое обсуждение разумно провести после объединения соответствующего материала.

1. ВВЕДЕНИЕ

В процессах, происходящих в живой природе, выделяются процессы, в ходе которых существенную роль играет обработка информации (о достаточно сложных описаниях явлений, ситуаций, структур и т.п., и, обычно, на основе той или иной организации памяти о прошлом) как основа для принятия (многократных последовательных) решений для достижения цели.

В [1] была намечена общая схема анализа таких процессов, названных там *информационными взаимодействиями*. Мы обсудим это понятие для живых систем — организмов весьма разного уровня: от бактерий и их предшественников до таких, которые обладают психикой и позже возникшим речевым интеллектом, развитой речью и, так называемым, “внутренним миром”. Попытка такого рассмотрения: от простого к сложному, от доорганизменного уровня к человеку с его сложным внутренним духовным миром — приводит к ряду особенностей изложения. Во-первых, невольно возникает как бы стиль эволюционного изложения; авторы должны определенно сказать, что они не стремились решать или даже обсуждать сложные и, может быть, мало разрешимые вопросы происхождения жизни, развития психики, речи, интеллекта и мышления. Во-вторых, авторы, не являясь, конечно, специалистами в этом длинном ряду соответствующих наук, не могли даже самым эскизным образом отмечать многие важные аспекты, не могли излагать вполне современные точки зрения (хорошо известно, что, например, в учебнике [2] многое менялось от издания к изданию, не говоря уже о периодике). Все, что мы хотели и могли — это отмечать в длинном ряду развития живого общие понятия именно с точки зрения информационного взаимодействия. Конечно, было бы желательно продолжить работу по уточнению, расширению и систематизации тех фактических сведений, на которых эти общие понятия основаны.

Возвращаясь к работе [1], заметим, что в живом: физический уровень — это уровень биомолекул, их комплексов, тканей и органов; сигнальный уровень присутствует как самостоятельно, так и в различных переходных формах от “чисто сигнального” ко все более сложным языкам (а затем и к иерархиям взаимосвязанных языков); лингвистический, семантический и прагматический уровни выступают здесь в их полной сложности как неразрывная триада синтаксис — семантика — прагматика (в семиотике и лингвистике обозначаемая словами семиотическая система).

2. ИСХОДНЫЕ ПОНЯТИЯ

1. Сигнал и информация. В анализе информационных взаимодействий фундаментальными понятиями являются понятия *сигнала* и *информационного сообщения* (“*информации*”). Эти понятия имеют синтаксический и семантический аспекты; первый состоит в анализе “внешней” формы, а второй — содержания сигнала или сообщения. Иными словами, в *первом случае* речь идет о некотором *конструктивном объекте*, т.е. записи (сигнала или информации); такая запись представляет собой слово (“фразу”, “текст”) в некотором алфавите или, иначе, — формулу (иногда даже “хорошо описанное” множество формул — теорию) в некотором логико-математическом языке. Синтаксическая сторона сигнала или информации тесно связана с представлением о передаче по каналу связи (с большим или меньшим учетом реальных свойств каналов, таких как: шум, помехоустойчивость, надежность, пропускная способность, кодирование — декодирование и т.п.). Во *втором случае* будем говорить о *семантической информации* или о *семантике* сигнала (сообщения). *Язык* (в самом общем смысле) определяется фиксированием его синтаксической стороны, т.е. некоторого набора слов, фраз, текстов, высказываний в соответствующем алфавите, а затем и в фиксировании семантики этого синтаксиса, т.е. семантики этих слов, фраз, текстов (иногда речь идет еще о фиксировании прагматики и некоторых других вспомогательных понятий, относящихся к языку). В частности, *прагматика* — это тот или иной способ придать определенный *контекст* некоторому классу текстов, высказываний. (Как известно, одно и то же высказывание в разных контекстах может иметь существенно разный смысл и побуждать, влечь существенно разные действия, следствия. Простейший пример из математики: аксиома о параллельных или континуум-гипотеза верна в одном контексте и не верна в другом. И это два существенно разных мира, или, как говорят, — две разных модели.)

2. Синтаксис и передача сигнала. Понятие количества (синтаксической) информации относится к записи, слову или формуле и определяется как длина этого слова (например, число битов в нем) или — другой вариант — как длина кратчайшей программы, алгоритмически порождающей это слово. Последнее представляет собой колмогоровское определение *сложности слова*. Заметим в этой связи, что теория колмогоровской сложности, таким образом, относится к чисто синтаксической стороне понятия информации.

Передача информации (т.е. некоторого слова) по каналу связи представляет собой цепочку дискретных возмущений, физическая природа которых не имеет теоретического значения. Передача характеризуется скоростью, надежностью, пропускной способностью и т.д. канала связи. Строго математическая и широко развитая теория информации (в общепринятом смысле этих слов) подчеркнуто считает своим предметом “информацию, отвлеченную от содержательной стороны дела”, т.е. от семантики сообщений. В ней, в сущности, речь идет об измерении статистической неоднородности при передаче возмущений по каналу связи. Обиходное или, например, медицинское употребление слова “информация”, конечно, связано с ее содержанием, смыслом, семантикой, т.е. со структурно организованными ситуациями, которые она описывает. Все это отнюдь не имеет стохастического характера и, вообще говоря, не связано с передачей по каналу. Будем различать эти две точки зрения как *статистическую (стохастическую)* и *семантическую*.

3. Канал связи. Упомянем общеизвестное описание стандартной модели *канала связи*. Имеются два “пункта” (или два базисных агента и тому подобное) P и Q . От P к Q поступает цепочка сигналов, которые принадлежат общему для них алфавиту A или двум разным алфавитам A_1 и A_2 (в этом случае подразумевается перекодирование в канале от A_1 к A_2). Далее, P и Q имеют каждый по своей семантике (часто совершенно различных), которые описывают некоторые классы ситуаций соответственно в “мирах, где живут” P и Q . Предполагается, что исходное слово у P является отражением некоторой “его потребности в его мире”, а полученное у Q слово оказывает или может со временем оказать влияние на мир Q (привести или способствовать определенным действиям со стороны Q в его мире).

Можно добавить, что как P , так и Q могут иметь каждый свои языки L_P и L_Q . Тогда P для передачи по каналу перекодирует слово своего языка в алфавит A_1 , а Q декодирует полученную цепочку сигналов в алфавите A_2 в слово его языка L_Q . Форму информационного взаимодействия, в которой более или менее присутствует канал связи (и посылка сигналов по нему), назовем *сигнальной*. Конечно, это лишь одна из таких форм.

4. Сигнальность. Отметим некоторые важные черты сигнальности, которые существенны для признания любого взаимодействия информационным.

1. Опосредованность. Воздействия P на Q происходят с участием промежуточных агентов (сигналов, неоднородностей), физическая природа которых безразлична для процессов в мирах у P и Q , о которых сообщается с помощью этих воздействий. Физическая реализация сигналов может быть заменена более или менее произвольно на другую; для этого нужно только обеспечить физическую перекодировку одного алфавита в другой. Сигнальное информационное взаимодействие не есть непосредственное взаимодействие, так сказать, базисных агентов (процессов) у P и Q через их прямое пространственно-временное или физико-химическое отношение. Опосредованность предполагает взаимодействие через промежуточных агентов, которые во времени и пространстве отделены от базисных, и не являются продуктами или субстратами базисных процессов. Словом, промежуточные агенты существуют независимо от базисных агентов и могут быть заменены другими промежуточными агентами. Во всяком случае, материальная природа вспомогательных агентов, сигналов, посредников не имеет значения. Например, в клетке сигнал часто влияет на направление той или иной реакции тем, что открывает (или активирует) хранилище, которое содержит вещество или энергию, или информацию (например, при обращении к ДНК). Это хранилище создается, наполняется и хранится независимо от базисного процесса, с которым его содержимое связано только, так сказать, логически; но высвобождение этого содержимого приводит (возможно, через систему усилителей) к изменению направления соответствующего биохимического процесса. В принципе такой механизм может соединять сигнал одного физического типа с процессами совсем другого физического типа. В лингвистике, семиотике этому соответствует принцип произвольности знака.

2. Предусмотренность. Хотя алфавит A , как сказано, может быть произвольно заменен другим, он должен находиться в согласии с языками L_P и L_Q . Это согласие должно быть обеспечено заранее, до того как произойдет обмен сигналами. Такая согласованность может быть достигнута естественной предысторией данного взаимодействия или прямой договоренностью “сторон” и т.д. *Предусмотренность* или *согласованность* означает, что при информационном взаимодействии хотя бы одна из двух участвующих во взаимодействии сторон до того была подготовлена к изготовлению или получению передаваемого сигнала, а также — к кодированию-декодированию и согласованию сигнала с основными процессами, идущими у P и Q . В более организованной форме предусмотренность означает договоренность, т.е. буквально договор между двумя сторонами, о форме и значении кодов (протоколов), о времени и канале передачи, и тому подобное. Предусмотренность включает в себя и согласованность (согласование) контекстов у принимающей и передающей сторон. Здесь должна участвовать память, в которой эти условия записаны, а также и объемлющая структура (“внешний мир”), отраженная в памяти и согласованная в важных чертах между P и Q .

3. Дискретность. В технике используют сигналы непрерывной природы, но все же и они подвергаются той или иной дискретизации (например, по теореме Котельникова). Во всяком случае в теории информации исходным понятием является дискретный и конечный алфавит. А.Н. Колмогоров писал “: в очень многих случаях изучение реальных явлений разумно вести, избегая промежуточного этапа их стилизации в духе представлений математики бесконечного и непрерывного, переходя прямо к дискретным моделям”. Изучение противоположности и взаимодействия непрерывного и дискретного будет специально обсуждаться дальше.

4. Выбор. Получение сообщения в мире у Q вызывает воздействие, которое в принципе изменяет некоторые протекающие у него процессы (сразу или потом) в одном или другом направлении из

некоторого конечного числа возможностей (случай бесконечного числа возможностей кажется скорее теоретическим); как минимум речь идет о двух возможностях. Сигнал соответствует выбору, который делает сам его отправитель или получатель, и/или отправитель заставляет сделать или способствует определенному выбору получателем. Часто выбор, который делается у P и Q , является выбором одного из их реальных состояний, которые могут по их последствиям жизненно важно различаться для P и Q . Этот выбор часто сводится к количественному различию в некоторой физической характеристике передаваемого сигнала. Связь этой характеристики с состояниями самих агентов не должна быть прямой; т.е. физико-химической частью процессов, определяющих их состояния, поэтому она должна быть “предусмотрена”, т.е. должен существовать дополнительный механизм влияния сигнала на эти процессы.

5. *Цепочечность.* Предполагается, что сообщение, передаваемое по каналу связи, является цепочкой сигналов. Это связано с дискретизацией времени, важной, например, для синхронизации взаимосвязанных процессов. Информацию можно передавать, например, картинками, но это более сложный способ, который все равно нуждается в последующей дискретизации. Мы будем рассматривать цепочки символов как основной случай. Синтаксис канала связи может иметь различный характер. Здесь можно отметить очевидную классификацию: каналов, передающие сигналы, которые прямо влияют на события в Q , или несущих слова специализированного языка, приспособленные к определенному классу событий в Q , или каналы, обладающие универсальными языками. Замечательным фактом является то, что такой универсальный язык возможен на основе простого цепочечного синтаксиса с передачей дискретных сообщений в дискретном времени.

Существенно, что уже алфавиты с очень небольшим числом символов (например, с 20 или 4, или даже 2) и соответствующие им всего лишь цепочечные конструктивные объекты (т.е. слова, последовательности символов) достаточны для организации фундаментальных биологических процессов. Хотя возможна кодировка всех слов в алфавите, содержащем даже только один символ, такой алфавит и такие слова, по-видимому, не могут служить в той же роли. (Уже два символа в алфавите дают экспоненциальное сокращение длин сообщений). Разумеется, чисто цепочечные алфавиты, обычно, осложнены “многомерным обслуживающим аппаратом”.

5. Семантика информации. Переходя к понятию “содержания (смысла) сигнала, сообщения или информации”, вернемся к обсуждению упомянутых выше и введенных для этой цели понятий “семантики” и “контекста”. Смысловая информация воспринимается в некотором контексте и имеет свое собственное содержание, которое мы будем называть семантикой сообщения. Семантика сообщения зависит или нет от его контекста. Эта семантика (в общем случае) имеет двойственный характер: то, *о чем* сообщение, и то, *для чего* оно используется, к чему оно побуждает.

Сообщение может более или менее широко варьироваться, сохраняя то же свое значение, т.е. давая одинаковое описание некоторого явления и в одинаковой мере служа достижению некоторой цели. При одном из возможных подходов говорят, что, в целом, семантика — это то, что позволяет считать некоторые формулы (фразы, тексты) “попарно эквивалентными друг другу, т.е. говорящими одно и то же относительно данного контекста”. Здесь можно различать два крайних случая: семантика “да-нет”, особенно характерная для изолированных сигналов, и семантика как перевод сообщения (формулы в некотором языке описания) в “картинку того, что было им (ею) описано”.

Такую “картинку” можно представлять себе как набор (частичных) структур, отношение между которыми организовано с помощью, например графа или дерева; т.е. как что-то подобное модели Крипке. Этот второй случай можно пояснить аналогией с характерным для логики представлением: от описания (от формулы, теории) к соответствующей структуре (модели).

Назначение семантики (во втором случае) состоит, прежде всего, в организации процесса многократного проигрывания (например, с целью построения плана действий для достижения фиксированной цели).

Семантическую информацию трудно оценивать с точки зрения количества. Здесь речь идет о степени ее достоверности, правдоподобности, которая оценивается числом (в случае, так называемой, нечеткой логики) или чаще элементом некоторого специально подбираемого частично упорядоченного множества (типичный пример которого — множество всех открытых подмножеств какого-то геометрического пространства).

Эти представления не расходятся с известной последовательностью в понимании самого термина “информатика”: сначала он относился к, так сказать, хорошо организованному библиотечному делу (независимо от носителя информации), затем и к синтаксической стороне сообщений, к ее передаче по каналу связи, теперь и к наиболее сложному аспекту информации — к ее содержанию.

В первую очередь, семантическая информация свойственна человеческой жизни, связывается с работой человеческого интеллекта, с приобретением и передачей знаний. Это понятие распространяется на общение животных, на сложные технические, человеко-машинные и социальные системы.

Многие процессы в живых системах уже на физиологическом и, тем более, на психическом уровнях имеют информационный характер. Например, несомненна информационность функционирования генетического кода или иммунной системы.

6. Семантика информации: о чем и для чего. Содержательная сторона информации находится с обеих сторон канала передачи (т.е. у P и у Q) и сводится к двум вопросам: *о чем и для чего* информирует процесс передачи информации, т.е. процесс прохождения сигналов по каналу связи.

В общих словах “о чем” означает: *о структуре* какого-то объекта или явления, — а поток сигналов содержит *описание* этой структуры. “Для чего” означает: для *управления* некоторым процессом или для *планирования* цепочки действий, имеющих *цель*, т.е. для составления плана действий по достижению цели, а получение данной информации способствует (успешно или нет) составлению и осуществлению этого плана. (Семантическая информация, например в теории А. Харкевича, соответствует цели в той мере, в какой она увеличивает вероятность ее достижения.) Управление и планирование направлено на построение (или на разрушение) функционирующей структуры, а с другой стороны, описание структуры нужно для ее целенаправленного построения или для составления плана действий. В развитом виде управление требует планирования, планирование требует модели процесса, а модель — описаний соответствующих структур. В случае конкретных информационных взаимодействий некоторые стороны этой схемы могут отсутствовать или быть представленными в зачаточном виде.

Для более подробного анализа этой схемы чуть ниже мы перейдем к обсуждению процессов, возникших в эволюции живого.

Если обозначить структуру, описываемую сообщением ϕ (т.е. семантику сообщения ϕ), у P как $\phi(P)$, а у Q как $\phi(Q)$, то естественно ожидать согласованности $\phi(P)$ и $\phi(Q)$, а именно, структура $\phi(Q)$ должна быть гомоморфным образом структуры $\phi(P)$.

Для создания описания ϕ (“создания информации”) само явление нужно сначала представить в огрубленном виде, “расчленив” и при том конструктивно: нужно в нем *усмотреть* его части и отношения между этими частями.

Заметим, что хорошим является язык с малым числом исходных отношений и операций, которых однако достаточно для того, чтобы составить формулы с богатой универсальной для своего класса явлений семантикой. Таков, например, язык теории множеств, который в одном из его вариантов имеет ровно одно исходное двуместное отношение “одно множество является элементом другого множества”, но при этом семантика формул этого языка охватывает всю математику и все точное естествознание.

Другим хорошим примером языка служит язык дуг, например тот, с помощью которого Ш. Губерман формализовал описание рукописных букв.

7. Влияния сообщения и память в канале связи. Обычно, полученное у Q сообщение ϕ не является совсем случайным для Q . Оно должно быть прочитано, т.е. понято в рамках его языка L_Q . Его содержание, обычно, должно оказать влияние на какую-то систему в мире у Q специально подготовленную для канала, по которому передано ϕ , а уже активность этой системы направит некоторый процесс по одному из возможных путей (т.е. действие сигнальной посылки опосредовано специальным приемником). Хорошей иллюстрацией служит действие электрического импульса на сокращение мышцы: оно опосредуется опорожнением кальциевых депо, благодаря которому приводится в действие актин-миозиновый комплекс, после чего кальций снова возвращается в депо. Также возможно, что пришедшее сообщение будет просто сохранено (“оставлено без последствий”), но тогда его нужно поместить в некоторую систему у Q , где оно займет “место, разумное с точки зрения семантики” языка L_Q . Таким образом, канал информационный связи может быть дополнен той или иной формой памяти у агентов P и Q .

8. Описание, управление, поведение. Итак, информация — это *описание* (на некотором языке) некоторой структуры. Она нужна для осуществления или оптимизации *управления* некоторым процессом или, во всяком случае, для воздействия на некоторое явление. Составление описания — это порождение информации. Цепочка последовательных и согласованных описаний, обычно, приводит к возникновению *модели* процесса или явления, к составлению плана (цепочки действий), ведущего к фиксированной *цели*. *Управление* — это целенаправленное воздействие для построения какой-либо структуры или для изменения поведения.

Важно определить границу, которая отделяет управление от “простого” воздействия. В первом приближении можно сказать, что граница там, где воздействие становится информационным, сигнальным. В простейшем случае такая структура может быть чисто геометрической или физико-химической (например, можно рассматривать чисто геометрическую симметрию кристалла или сверх того учитывать его химический состав), статической или динамической (т.е. структурой процесса). Она становится функциональной, если является составной частью более широкой структуры; функционирующей, если она поддерживает более широкий динамический процесс, т.е. поддерживает упорядоченное изменение во времени некоторой системы. Структуры некоторых процессов и явлений основаны на своих внутренних языках (например, их имеют основные процессы в клетке). В этом случае язык описания более высокого уровня оказывается как бы метаязыком по отношению к языку нижнего уровня; и можно говорить о хорошей (или плохой) согласованности таких двух языков.

9. Интеллектуальность сигнальности. Вернемся к общему представлению о сигнальном, информационном воздействии, например, со стороны среды обитания и в связи с задачей достижения цели.

Здесь мы вынуждены привлечь ряд понятий. Система (или, как иногда говорят, интеллектуальная система) в дискретном времени получает входную информацию T_n , которая активизирует (актуализирует) “относящуюся к ней область памяти”; в результате чего образуется текущий контекст (ситуации) T_n . Кроме того, T_n зависит от цели (мотива) φ , которую (который) система имеет в данный период времени (шкала времени с переменной n меньше, чем этот период времени) и от “мира”, в рамках которого находится система (в смысле: “мир семьи”, “мир работы” и т.д.; выбор этого мира зависит и от T_n). Система реагирует на T_n , стремясь достичь цель φ в контексте T_n ; цепочку последовательных реакций можно представить себе как стремление системы построить структуру, в которой выполняется φ , интерпретируя язык (или, что то же самое, символы) входной информации. Подробнее это будет обсуждаться во второй части статьи.

Заметим, что здесь присутствует переход от физически данного явления к его “отражению” в некоторой системе, имеющей, тем самым, интеллектуальные черты (способной к описанию, моделированию, к целенаправленному управлению, планированию). Зародыш такой интеллектуальности присутствует уже в сигнальности. Формы сигнальности достаточно разнообразны, здесь также требуется разбор частных случаев сигнального взаимодействия.

Информационное воздействие (сообщение информации) от A к B означает, что B получает стимул к выбору состояния или действия, при этом стимул, вообще говоря, не вынуждающий B перейти в это состояние или совершить это действие, но в той или иной степени (так сказать, “от 0 до 1”) влияющий на его совершение в подходящей ситуации текущей или будущей. Это и означает, что информационное воздействие является “опосредованным”, т.е. воздействием, не вызывающим непосредственного изменения в каком-то из основных процессов в B ; а является воздействием на специальную часть C в B , которая имеет независимое от основных процессов в B существование, но изменения в которой могут вызвать соответствующие изменения в них уже независимо от дальнейшего воздействия со стороны A . Заметим, что другие агенты могут влиять на ход событий в B , оказывая воздействия на эту связь A и B , т.е. на C ; например, изменяя порог воздействий C на B .

10. Рубеж психики. Появление психики разделяет эволюцию на два принципиально различных этапа. Основой допсихической жизни как отдельного организма, так и всего живого, является построение организма, построение его функционирующих структур (в совокупности составляющих организм), а также копирование организмов, но взаимоотношение этих организмов с окружающим миром не имеет характера целенаправленного управляемого поведения. Хотя поведение строится в значительной мере уже на сигнальной основе, оно в большей мере управляется средой, чем внутренними потребностями и внутренними “решениями”. Появление психики приводит к тому, что внешнее поведение организма становится определяющим фактором его выживания. Психика автономизируется и из “просто органа” с функцией управления поведением становится самим организмом, становится постепенно “Я” организма, а сам организм превращается во “внутреннюю среду” этого “Я”, в значительной мере “равноправную с внешней средой”. Указать точную границу возникновения психики в ряду известных видов очень трудно, если только возможно. Это тем более сложно, что сейчас невозможно указать и точный критерий наличия психики.

11. Самопроизвольность и информационное управление. В живых организмах многие существенные процессы протекают без управляющих воздействий вследствие заранее созданных общих физико-химических условий, статистической динамики, термодинамики. Так, секрецию, управляемую гормонами, отличают от “конститутивной”; многие большие молекулы, например микротрубочки, образуются самосборкой и т.п. Информационное управление есть воздействие, опосредованное сигналами, информационными сообщениями. Иными словами, роль информационного управления заключается в изменении условий (контекста), в которых протекает процесс. Если, скажем, это изменение температуры, требующееся для кристаллизации, то управлением будет не прямое воздействие на движение молекул (нагреванием или охлаждением), а воздействие через сигнал, вызывающий нужное изменение температуры.

Заметим, что в теории обратной связи также различают обратную связь, непосредственную по выходу, или опосредованную сигналами измеряющих приборов.

Самопроизвольное образование больших молекул есть процесс, хоть и не сигнальный, но все же в некоторой степени языковой. Здесь существенно наличие своего рода языка больших молекул: элементы соединяются под действием физико-химических сил, но при этом важно, что они имеют определенную форму, имеют концы притяжения и отталкивания и т.д., так что их соединение происходит по определенным разрешающим и запрещающим правилам. Уже здесь происходит взаимодействие логики и статистической кинетики (термодинамики), что подробнее обсуждается в пунктах 8 и 9 раздела 2.

3. НЕКОТОРЫЕ ОБЩИЕ ПРИНЦИПЫ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ В ЖИВОМ

1. Эволюционный подход. Конечно, центральное положение в естественной истории информативности занимает человеческий интеллект. В понимании его строения и функционирования имеются два аспекта: эволюционный, основанный на анализе роли информативности на различных этапах

эволюции, и психолого-лингвистический, основанный на анализе речевой и психологической функций интеллекта. Заметим, что в обоих аспектах в качестве математического инструмента естественно использование аппарата математической логики и теории алгоритмов.

Здесь наш подход к анализу понятия информационности будет эволюционным. Мы рассмотрим, конечно, очень условно и конспективно вопрос о том, как информационность возникает и развивается в естественных условиях.

Эволюция информационности является важнейшей частью общебиологической эволюции; которой мы, таким образом, будем касаться. В третьей части этой работы коснемся развития собственно человеческого интеллекта.

2. Об эволюции информационности. В искусственной системе предусмотренность сигнальной связи является результатом обдуманного и спланированного построения. В естественных информационных системах роль планирующего органа берет на себя, в том числе, эволюция. “Тончайше отрегулированные, согласованные системы сложились в результате длительной эволюции” — обычное выражение в биологической литературе. Эта согласованность имеет разнообразные формы, одна из которых состоит в возникновении цепочек действий, процессов, которые служат как бы фиксированными реализациями важнейших алгоритмов. Например, метаболические пути, канализация морфогенеза, цепные рефлексы и т.д.

Такие цепочки процессов часто соседствуют с циклически организованными процессами: циклы клеточного деления, реакции дыхания и пищеварения, поведенческие циклы питания, воспроизводства и т.д. В клетке основные системы состоят из большого числа таких иерархически организованных цепочек и циклов.

Здесь “согласованность” также включает и согласованность соответствующих контекстов.

3. Предшественники устойчивости и оптимальности. В математическом плане естественно упомянуть об аттракторах динамических систем, их стремлении к равновесному состоянию. Кроме состояний равновесия известны еще устойчивые систематически повторяющиеся цепочки процессов, в частности, циклов, связанных с глобальным развитием и глобальными циклами звезд, планетных систем и т.д. В этом случае имеется “как бы предусмотренность” и “как бы выбор”. В стремлении повторяющихся процессов ко все большей определенности (например, для планетной системы — небольшое число планет с определенными соотношениями их параметров) можно усмотреть аналогию с “проложенными” эволюцией метаболическими путями и креодами морфогенеза, но ясно их отличие. В добиологическом случае нет внутреннего языка, нет воспроизводства-копирования действующих лиц, нет сигнальности. На “чисто” физическом и на биологическом уровнях цепочки процессов складываются на основе более или менее известных принципов оптимальности.

4. Дискретность. Базисным свойством “косной” природы, на основе которого строится информационная составляющая жизненных процессов, является как дискретность, так и (на основе дискретности) постепенное усложнение геометрической формы единиц строения материи. Дискретность является фундаментальным свойством материи и ее обнаружение в конкретных областях науки оказывает большое влияние на соответствующее направление: открытие закона кратных отношений улучшило понимание химического строения вещества, генетика приблизила биологию к статусу точной науки. Другой пример, построение видовых филогенетических деревьев на основе согласования белковых филогенетических деревьев создало в эволюционной теории важное направление, обладающую большей предсказательной силой (что, конечно, не заменило и других методов изучения эволюции).

Значение дискретности, т.е. относительной неизменности элементов, состоит в том, что достаточно малые влияния оказываются нулевыми, устойчивость системы существенно повышается (что иллюстрируется преимуществами цифровой техники над аналоговой) и т.д. Правда, при этом ма-

лая ошибка (не погрешность, а именно ошибка) может повлечь суровые последствия (что видно на примере цифровой техники и генетики).

Значение дискретности состоит еще в том, что она позволяет вести статистические подсчеты на более точном основании, поскольку в этом случае имеется ясное представление о пространстве элементарных событий, и нет необходимости вводить априорные “признаки”. (Однако, увлечение статистическими подсчетами может приводить к бессмысленным результатам, если не учитывать, что элементарные события далеко не всегда независимы.) Кроме дискретности элементов (например, мутации происходят случайно и в отдельных элементах) также важна *логическая дискретность*, т.е. составленность явлений из частей и их отношений (операций над ними), которая дает, так сказать, логический каркас явления.

Конечно, дискретность в определенном смысле соседствует с непрерывностью (например, в концепции частица — волна). Дискретность элементов часто сочетается с непрерывностью соответствующего аппарата, например, при термодинамическом описания процессов, идущих с такими элементами. Непрерывность и вообще может быть выражена на дискретном языке, как показывает, например, комбинаторная топология относительно теории многообразий; или когда в топологии рассматриваются конечные топологические пространства.

В этом смысле биологическую эволюцию можно считать непрерывно развивающимся во времени процессом, хотя она и заключается в разветвленной цепочке дискретных шагов по передаче информации от организма к организму.

5. Языки. В добиологической природе имеются сложно организованные, но не информационные системы также образованные “в результате длительной эволюции”. Например, таков климат Земли, устойчивость которого поддерживается ледниками, Солнцем, океанами, вращением Земли и т.п.

Отличие (имеющее информационное значение) биологической системы, “складывающейся в эволюции”, от добиологической системы, в частности, состоит в том, что в первой из них развивается язык (точнее, система языков), и сама эволюция происходит на этой языковой основе. В частности, сигнальность эволюционирует в языковые системы и, в конечном счете, приводит к естественному человеческому языку. А затем — и к компьютерным, логико-математическим языкам, что не мешает последним быть в некоторых отношениях неизмеримо проще, чем естественные сигнальные и языковые системы, возникшие в эволюции. Можно говорить о семиотической, языковой эволюции как о составной части в общей эволюции информативности. Естественные языки обычно имеют командную и эмоциональную составляющие.

6. Добиологическая природа языковых систем. В добиологической природе можно увидеть предпосылки для возникновения языков уже в небольшом числе стабильных элементарных частиц и отсюда в небольшом числе стабильных атомов, существенных для биохимических процессов, в небольшом числе (немного более сотни) химических элементов, на основе которых строится существенно большее число “фраз-молекул” по определенным правилам (в основном ковалентной связи). Затем появляется возможность построения неограниченно большого числа органических полимеров. Вступает в действие биологическая эволюция: возникают мембраны и в ограниченных ими частях пространства, клетках происходит строгий отбор небольшого числа (сравнительно с возможным) реально используемых органических молекул с особыми свойствами и связанных с ними биохимических циклов. Уже на основе этого синтаксиса возникают сигнальность, семантика и контекст, т.е. язык, соответствующий этому синтаксису.

7. Алгоритмическая природа биологической эволюции. Сильно утрируя, биологическую эволюцию можно назвать (в отличие от добиологической, скажем, звездной) алгоритмической, представляя ее себе как постепенное увеличение запаса фраз, усложнение самих фраз и их семантики, прагматики, т.е. в постепенно усложнение языка эволюции. Этот язык лежит в ее основе и служит как бы ее костяком. Благодаря ему, невероятное разнообразие организмов может быть описано компакт-

ным образом, а само разнообразие возможно только на основе общего глубинного (биохимического) языка, представляясь как бы (с известными оговорками) разнообразием фраз, которые можно построить из данных элементов по данным правилам при постепенном усложнении самих элементов, самих правил, их семантик и контекстов.

Более конкретно, биологический мир представляет собой единую систему: вероятно, биомолекулы из организма одного вида по трофической цепочке могут попасть в организм другого достаточно далекого вида, не разрушаясь до неорганического уровня и быть там усвоенными с помощью общих биохимических механизмов; горизонтальный перенос генов, например с помощью плазмид, из одного генома в другой — еще один пример.

8. Взаимоотношение “алгоритмичности” и “статистичности”. Благодаря такому характеру эволюции она происходит дискретными шагами и каждый шаг представляет собой выбор из сравнительно небольшого числа возможностей, описанных на подготовленном к этому моменту языке. Этот выбор происходит “естественным путем”: каждый шаг осуществляется в результате “вероятностного розыгрыша”.

Взаимоотношение “алгоритмичности” и “статистичности (стохастичности)” является характерной чертой биологической эволюции (до известной степени, и технической эволюции). Несмотря на то, что каждый эволюционный выбор как бы увеличивает определенность (и, тем самым, уменьшает энтропию), все же на текущем уровне возникающей системы сохраняется “определенная неопределенность”, допускающая дальнейшую эволюцию на основе статистического выбора.

По этому поводу у Колмогорова имеется выражение “остаточная энтропия”, относящееся к владению поэтами словесным материалом (насколько велико количество слов, остающихся для выбора очередного слова в нужное место стиха после того, как произошло ограничение всеми до того выбранными словами и условностями стиха).

Важно, что в каждый момент имеется небольшое число вариантов дальнейшего развития, определяемых уже построенной системой и выразимых на уже построенном к этому моменту языке (закон гомологических рядов, по-видимому, может пониматься в этих рамках). В то же время имеется “вероятностное распределение” возможных выборов для следующего шага в построении системы — распределение, определяемое соответствием таких выборов предыдущим шагам, внешним условиям и характеру устойчивости получающегося результата.

С размножением единиц энтропия возрастает опять, но это происходит уже на новом уровне, на основе вновь появившихся структур.

Таким образом, представляется, что изучение сочетания формально-логического развития системы конструктивных объектов (управляемого логическим формализмом) со стохастической динамикой, ограниченной этим формализмом, является фундаментальной задачей. Иными словами, если на уровне человеческого интеллекта предусмотренность и выбор могут иметь вид произвольного соглашения, то на биологическом уровне они носят более природный характер и определяются эволюцией, которая сочетает логику и стохастику.

9. О модели, сочетающей алгоритмичность и стохастику. Конечно, стохастическая часть внутриклеточных взаимодействий также является набором алгоритмов, по крайней мере, на компьютерном уровне. Поэтому речь идет о взаимоотношении двух классов алгоритмов, идущих от генома и от химической кинетики. Условно представим себе, что это классы (как принято говорить в математической логике) “экспоненциальных” и “полиномиальных” алгоритмов.

Итак, “термодинамический котел” — это собрание полиномиальных алгоритмов с пустыми слотами для подстановки в них оракулов — экспоненциальных алгоритмов, которые собраны в некотором хранилище-“геноме”. Без заполнения всех слотов полиномиальный алгоритм работает очень медленно (практически не работает). Вбрасывание вектора оракулов (набора “ферментов”, соответствующих какой-то одной цепи реакций) из генома приводит к реальной, быстрой работе соответствующих

полиномиальных алгоритмов; в результате они производят экспоненциальной сложности объекты, которые затем могут служить входами тех же полиномиальных алгоритмов. До их производства входами полиномиальных алгоритмов служили только полиномиальные объекты.

Обратная связь состоит в том, что так произведенные экспоненциальные объекты присоединяются к геному и, тем самым, останавливают выпуск одних оракулов или сигнализируют о необходимости выпуска других.

В некотором адекватном темпе в котле происходит разрушение экспоненциальных объектов; полиномиальные как объекты, так и алгоритмы, не разрушаются.

Сверхзадача этой ситуации разложить, вывести некоторые экспоненциальные объекты из котла, ввести некоторые такого рода объекты извне и накопить определенный избыток определенных экспоненциальных объектов. (Конечно, мы следим не за одной клеткой, а за их коллективом и статистически, т.е. тут присутствует “внешняя цель”.)

Сам геном случайно меняется, но не равномерно: некоторые места не так важны, другие более защищены. Это приводит к невозможности производства одних экспоненциальных объектов и к появлению новых другого типа таких объектов (невозможность накопления одних веществ и появление новых свойств, например устойчивости к пенициллину). Такого рода описания (“переходные” от чисто биологических к “полуматематическим”), на наш взгляд, представляют интерес на данной стадии вопросов.

Теперь поставим вопрос: что в добиологическом мире является предшественником этих черт, основой, из которой они “возникают”.

4. ДОКЛЕТОЧНЫЕ УРОВНИ РАЗВИТИЯ

10. Развитие геометрической формы единиц материи. Возвращаясь к естественной предыстории информативности, кратко упомянем развитие геометрической формы единиц материи как важный процесс, приведший к возникновению живой материи и вместе с ней информационной составляющей естественных процессов. Как и в других пунктах это описание делается очень эскизно, в сущности, на наивно наглядном уровне.

Непосредственное участие в биохимических процессах принимают три частицы: электрон, протон и фотон (отдельную роль играет нейтрон), которые, наивно говоря, представляются имеющими точечную симметрию. Уже атомы с их электронными орбиталями нарушают эту симметрию, в химии атому приписывается определенное геометрическое строение в форме ежа — набора векторов, выходящих из общей вершины под фиксированными углами друг к другу.

Это еще не конструктивные объекты в смысле отсутствия у них регулярной составленности из знаков фиксированного алфавита по фиксированным правилам. Действительно, атомное ядро (не говоря о более элементарных частицах) не состоит из своих частей (протонов и нейтронов) в строгом смысле. Для конструктивного объекта его части, по меньшей мере, не должны исчезать в объединении, должны быть различимы в целом (хотя бы в деформированном виде). Элементарные частицы при объединении не *составляют* новую частицу, они исчезают и рождается новая частица как волна, возникшая из двух встретившихся волн.

11. Синтаксис молекул. Углы, характеризующие геометрическую форму атомов, в значительной мере определяют синтаксис первичного химического языка — структурные формулы молекул. Это — язык графов с фиксированными величинами углов и длинами ребер, имеющий очень сложный синтаксис (сложные правила порождения новых молекул). В такой ситуации трудно приписать формулам адекватную семантику. В таком синтаксисе трудно породить по фиксированным простым правилам регулярные конструктивные объекты — структурные формулы молекул.

Если число элементарных частиц — участников биохимических процессов ограничено четырьмя, а число химических элементов (таблицы Менделеева) немногим более сотни и имеется физически

определенная граница для номеров химических элементов больших, скажем, 150, то число неорганических веществ существенно больше, вероятно, несколько тысяч. Однако и здесь можно представить себе, что практически неограниченное разнообразие таких молекул невозможно.

В пределах этого химического языка в процессе добиологической химической эволюции произошел отбор небольшого числа молекул с небольшим числом правил их соединения, который создал более простые цепочечные языки на основе полимеризации, легшие в основу биохимии клетки. Шаги этой эволюции были определены неизвестными нам начальными условиями и случайностью. В основе своей химия клеточных процессов это — химия небольшого числа химических элементов: шести элементов С, О, Н, N, P, S и воды. Имеется еще десяток элементов, которые играют важную роль в отдельных процессах (некоторые в микроколичествах).

Химический состав живого вещества в основном един для всех организмов, в некоторых чертах он связан с составом земной коры: например, плазма крови по составу солей близка к морской воде, — но в целом существенно отличается от нее.

Итак, простота цепочечного синтаксиса позволила дальнейшие усложнения геометрической формы биомолекул, например, их гибкость, их локально меняющуюся форму, их сворачиваемость в определенные конфигурации (вторичная и третичная структуры) и так до образования различных конформаций и, например, аттенуаторной регуляции.

Основные биомолекулы принадлежат к четырем типам: полисахариды, липиды, белки и нуклеиновые кислоты. Эти большие молекулы построены из элементарных частей, являются конструктивными объектами. Однако только белки и нуклеиновые кислоты имеют непосредственное информационное значение для клеточной динамики. В частности, в той связи, что синтаксис полисахаридов и липидов не цепочечный, они допускают, особенно первые, достаточно сложные ветвления и циклы; а с другой стороны, белковые и нуклеиновые цепочки построены с помощью алфавитов, содержащих соответственно (примерно) 20 и 4 буквы с удивительно простой семантикой.

Вероятно, элементы этих алфавитов могут располагаться в цепочки совершенно произвольно, т.е. возможности конструирования белковых молекул в принципе неограниченны. Если и существуют ограничения, скажем, на длину белковой цепи, то число безусловно возможных конструкций столь велико, что подавляющее большинство возможных комбинаций просто никогда не реализовывалось.

12. Функция формы молекул (начальная семантика языка цепных молекул). Эти цепные молекулы-белки существуют обычно в свернутом виде (вторичная и третичная структуры). Расположение спиралей, возможные альтернативы были найдены отбором и имеют важное функциональное значение; например, вспомним об аттенуаторной регуляции генов и оперонов. В биологических системах форма молекул играют решающую роль в механизмах, с помощью которых регулируется экспрессия генов и поведение хромосом во время удвоения клетки, скорость мутаций, репарация повреждений, вырезание интронов и метилирование, и т.д.

Форма молекул играет особенно важную роль в ферментативной функции белков. Это катализ нового типа, основанный на геометрии и информативности, на удержании реагентов, с которыми фермент вступает не в прочную ковалентную связь, а в множественные слабые связи (водородные, ионные, ван дер Ваальсовы); а также — на образовании сайтов узнавания лигандов и т.п. Слабость связей позволяет в огромное число раз (в миллионы) ускорить катализ по сравнению с неорганическим. С другой стороны, регуляторные и ферментативные функции молекул требуют множественности связей, и в результате соединение молекул приобретает тип “ключ-замок” (усложненная комплементарность).

13. Возникновение органических молекул и “первичного бульона”. Этап химической эволюции. По некоторым представлениям в первичной атмосфере Земли почти отсутствовал кислород, но было много углекислого газа, аммиака, паров воды, водорода, азота, метана и др. Из этих веществ под действием ультрафиолетового излучения и грозových разрядов могли образовываться достаточно

сложные органические молекулы, в частности, аминокислоты и нуклеотиды. При остывании Земли вода смогла конденсироваться на ее поверхности в виде жидкости, образовав моря и океаны. В них растворялись органические молекулы, а углеводороды плавали по поверхности, образуя пленки. Так возник “первичный бульон” (по выражению Дж. Холдейна).

Согласно представлениям И. Пригожина в открытых системах имеется поток веществ и возникают диссипативные структуры. В таких структурах и могла идти химическая эволюция и возникать самоорганизация. Основным механизмом, обеспечивающим самоорганизацию в открытых системах, является естественный отбор. Конкретные модели такой химической эволюции в открытых системах были предложены московским химиком А.П. Руденко (1969) и немецким химиком М. Эйгеном (1971). Они предложили теории химической эволюции открытых каталитических систем и отбора на основе совершенствования и дифференциации в них каталитических реакций. В качестве особого класса самоорганизующихся химических цепей Эйген выделил гиперциклы, которые возникают, когда образуется цепочка связей между элементарными каталитическими циклами. Предполагается, что химическая эволюция шла в особых структурах. Согласно Опарину это коацерваты — особые структуры, образующихся в растворах гидрофильных коллоидов. Коацерваты накапливают вещества из окружающего раствора (при этом интенсивность химических процессов в них возрастает), увеличиваются в размерах и распадаются на более мелкие, т.е. в некотором смысле размножаются. Предлагаются и другие варианты таких структур. Каковы бы ни были эти структуры, они представляли собой открытые системы, в которых происходили химические реакции и процессы совершенствования каталитических систем на стадии химической эволюции. В них и могли постепенно формироваться системы обмена веществ и наследственности.

Сейчас предполагается, что важнейшую роль на этапе первичной эволюции могли играть молекулы РНК. В 1982 году было обнаружено, что некоторые рибонуклеиновые кислоты, как и белки, могут обладать ферментативной активностью — такие РНК называют *рибозимами*. Некоторые рибозимные РНК могут “сшивать” аминокислоты в белки. Если эта гипотеза верна, то первичным источником наследственной информации были РНК. В ходе дальнейшей эволюции эта функция перешла к ДНК. Дело в том, что РНК (как и белки) склонны сворачиваться в разные структуры, с которых труднее считывать информацию, а ДНК образуют двойную спираль. С другой стороны, среди тех белков, которые возникли в результате каталитической активности РНК могли оказаться и такие, которые были более сильными катализаторами тех или иных реакций, чем сама РНК. В результате наследственная информация стала храниться в ДНК, роль ферментов стали выполнять белки, а РНК стала выполнять иные функции: в основном — роль посредника между ДНК и белками.

14. Мембраны и их роль. Еще одна составная часть предбиологической информативности — появление липидных мембран. Возможно, благодаря этому событию возникли первые протоорганизмы (названные Опариным коацерватами), которые отделились (изолировались) от остальной химической среды. Поэтому химическая эволюция получила возможность независимых экспериментов. В таких коацерватах могли отрабатываться фундаментальные клеточные механизмы: цикл деления, клеточная генная память, направленный обмен веществ с внешней средой — таким образом возникала логика клеточного функционирования.

Уникальные свойства липидных мембран (например, полупроницаемость) определяются, прежде всего, двумерной упорядоченностью (которая позволяет считать их кристаллическими образованиями) и достаточно свободной подвижностью молекул в них (так что их можно называть двумерными жидкостями).

В этой двумерной жидкости плавают большие молекулы (белки и гликопротеины), которые в клетках служат для разных целей и, в том числе, для информатических: например, в качестве рецепторов и каналов для прохождения сигнальных молекул.

Важное значение имеет также замена трехмерной диффузии двумерной, благодаря чему существенно повышается вероятность встречи реагентов и значительно ускоряются реакции.

Возможно, что липидные мембраны появились как пленки на поверхности водоемов. В абиогенных условиях мембраны могли образоваться (особенно при наличии углеводородных цепей с гидрофильным и гидрофобным концами) в водной среде с полимерами двух типов, которые плохо взаимодействуют друг с другом и хорошо — с водой. Тогда возможно появление фазовых границ и пузырьков-коацерватов и соответствующее протекание в них основных ферментативных реакций.

Итак, важнейшая роль мембран состояла в вычленении дискретных единиц из непрерывной среды, способных к активному поддержанию своей целостности.

15. Размножение. Основной процесс, обеспечиваемый молекулами-переносчиками наследственной информации, — размножение, которое сначала идет на молекулярном уровне (и которое первично осуществлялось без участия большинства из ныне обычных молекул клетки).

Воспроизведение — сторона биологической эволюции, существенная, в частности, в связи с “остаточной энтропией”. Жизнь организма это цикл, который повторяется в других организмах. Накопленные полезные изменения должны передаваться по возможности без изменения, поэтому необходим механизм копирования и в виду неизбежных ошибок, необходимых эволюции, — многократного копирования. Оно служит одним из механизмов, предохраняющих от вырождения: вновь сделанный выбор, чтобы сохраниться, должен быть размножен и в то же время размножение позволяет проверить сделанный выбор, а также и обеспечивает “остаточную энтропию”: новые выборы на основе сочетания логики и стохастики.

16. Синтез молекул. В биологических системах молекулы белков и нуклеиновых кислот строятся, они не могут возникать спонтанно (во всяком случае с произвольной последовательностью элементов цепи). Такое построение осуществляется для белков и нуклеиновых кислот сложной системой передачи информации от генов на основе развития древнего механизма комплементарности. Такое построение заключается в последовательном присоединении аминокислот к растущей последовательности, задаваемой молекулой мРНК, считанной с соответствующего гена при участии рибосомы и комплекса, включающего транспортные РНК. А также — во взаимно однозначном копировании двойственной цепи (осложненном различными вторичными механизмами, включая РНК-полимеразу).

17. Примитивный организм. Представляет интерес вопрос о “минимальном размере” возможного “примитивного организма”: размер ДНК, число белков, основные типы биомолекул, их взаимодействия и назначения. Известны такого сорта оценки, например число дальтонов — приблизительно массой протона. Можно думать, что организменному уровню предшествовал переходной доорганизменный, преджизненный этап эволюции, когда указанные выше характерные процессы и информационные черты еще не были отлажены, а вместо организмов имелись сгустки химических реакций, уже отделенных мембраной от внешней среды: внутри которых происходили процессы, кристаллизовавшиеся, в конечном счете, в налаженные процессы внутриклеточной динамики. Первые ответы на эти вопросы были даны в работах 60-х годов В.М. Андреева, который указал, что для минимального метаболизма автономной самовоспроизводящейся клетки необходимо иметь не менее 118 типов белков, что отвечает примерно 100 млн. D ДНК.

5. КЛЕТочный уровень развития

1. Клетка и начало биологической эволюции. Клетка это организм, единица жизни, а с другой стороны, это естественное продолжение эволюции единиц материи, возникшее в “результате” стремления к состоянию упорядоченного равновесия (внутреннего и с окружающей средой) хаотической химической динамики в изолированных областях “мирового химического бульона”.

С момента появления клетки начинается биологическая эволюция, в том числе дарвиновская. Суть последней — в силу ограниченности доступных средств существования, выживают те типы организмов, которые оставляют большее число своих копий и, во всяком случае, не меньшее, чем число погибающих копий. Наследуемые малые изменения (т.е. дискретная непрерывность) приводят

к постепенному преобразованию одних биологических видов в другие, все более приспособленные к данным жизненным условиям (в которые входят не только физико-химические свойства среды, но и биология среды, включая генную среду окружающих видов).

2. Соотношение случайности и логики в эволюции. Снова отметим взаимовлияние статистических законов случайности и развития логического строения организмов и их сообществ. Эволюция приводит к построению чрезвычайно разнообразных организмов в результате дискретной языковой основы строения самого организма и допустимых дискретных изменений, которые накапливаются благодаря отбору и биологической памяти, языковой эволюции. Малые изменения случайны в том смысле, что они не имеют определенной направленности в какую-то сторону (не способствуют улучшению или усложнению сами по себе), не “задумываются” об их эволюционных последствиях. Они одновременно не случайны в том смысле, что мутации происходят не во всех местах генома равновероятно (можно вспомнить закон гомологических рядов Н.И. Вавилова), скорость мутаций может регулироваться и, главное, эволюционный отбор действует лишь на изменение того, что уже имеется.

3. Внешняя среда направляет эволюцию. Эволюционный процесс направляется изменением внешних условий. Его информационная составляющая заключается в жесткой языковой логике генома, направляющего построение организма, и биологической памяти наследования малых изменений, соответствующих условиям среды. В зависимости от текущего взаимоотношения генома и внешних условий, изменения могут вести к стабилизации (стабилизирующий отбор), быть подчиненными некоторому тренду (направленный отбор) или вести к расщеплению популяций (дисруптивный отбор). Во всех случаях ведущую роль играет внешняя среда, динамика генома *реактивна*, обладая только случайной мутационной активностью.

В понятие среды для данного гена и для систем генов входит также и внутренняя геномная среда: гены отбираются на согласованность с внешней для организма средой через их кооперативную согласованность.

4. Конструктивная сложность в эволюции. Путь эволюции это — не только усложнение организмов (виды могут становиться и более примитивными). Генетическая память сохраняет возможные вариации черт организмов и тем *расширяет* запас возможных реакций на изменения внешней среды, эта генетическая память оказывает существенное влияние на эволюцию, определяет ее направление в сторону приспособленности и усложнения. Возникающая сложность и логика генома отражают сложность и логику среды организма, и в этом смысле геном несет все возрастающую информацию о внешнем мире.

Язык генома является языком не только в силу своего цепочечного синтаксиса, осложненного различными генными механизмами. Он имеет также многоуровневую семантику, выражающую как непосредственную его функцию — построение белковых молекул, — так и более отдаленные: построение организма и его поведения и в конечном счете отражение (“выражение”) его биологической среды. Языковая система генома имеет также и прагматику: функционирование гена зависит от “контекста” — от актуального состояния организма и среды, особенно от стадий эмбриогенеза и проч.

5. Исправление ошибок копирования в клетке. При репликации ДНК возникают ошибки. Они могут быть обусловлены дефектностью молекул фермента репликации (ДНК-полимераза) и тем, что в момент репликации в клетке оказалась какая-то активная молекула, влияющая на репликацию, и генными мутациями, которые вызываются мутагенами, и т.д. Казалось бы из-за совместного действия всех этих факторов генные мутации должны иметь большое влияние на наследственность. Однако, в клетках имеются системы репарации, которые следят за сохранностью информации, хранящейся в ДНК. Сейчас у кишечной палочки известно более 50 генов, контролирующих процессы репарации. Эти гены кодируют ферменты, которые умеют, например, вырезать поврежденные участки одной цепи ДНК. ДНК-полимераза достраивает это место в цепи до нормы, а ДНК-лигазы “зашивают” образовавшийся разрыв в месте встроеного участка. Такие же системы обнаружены у растений и

животных. Имеются специальные ферменты, которые устраняют повреждения ДНК, создаваемые ультрафиолетом. Системы репарации устраняют разрывы, возникающие в цепях ДНК под действием рентгеновских лучей и других излучений.

У всех клеточных организмов роль наследственного материала играет двухцепочечная молекула ДНК. Возможно, одна из существенных причин, почему именно ДНК в ходе эволюции стала играть особую роль состоит в том, что наличие второй (“запасной”) цепи позволяет исправлять дефекты, возникающие в одной из цепей. Таким образом, наличие избыточной информации (во второй цепи) увеличивает надежность хранения информации. Таким образом, подобно тому, как на уровне организма имеются системы исправления повреждений (заживления ран, а у некоторых организмов и регенерации утраченных частей тела), на молекулярном уровне имеются системы исправления повреждений наследственного материала.

6. ДОПСИХИЧЕСКИЙ УРОВЕНЬ РАЗВИТИЯ

Рассмотрим теперь информационные аспекты организменного уровня. Начиная с появления клеточной формы жизни и до возникновения психики можно отметить несколько критических событий, из которых важнейшими являются появление *эукариотов* и затем *многоклеточных*.

7. Прокариоты (бактерии). Основную биомассу на Земле составляют бактерии. Они достаточно далеко ушли от первичных клеток, и все же именно они являются их прямыми наследниками. Возникновение первых клеток относят примерно ко времени в 3.8 млрд. лет назад.

Особенность бактерий состоит в их гораздо большей, чем у других организмов, способности химически взаимодействовать с неорганическим окружением. Они, в частности, окисляют некоторые соединения серы и азота, синтезируют метан из CO_2 и H_2 и аминокислоты из простых углеводов. Кислород атмосферы — продукт деятельности сине-зеленых водорослей океана (как, впрочем, и зеленых растений). Азот атмосферы фиксируется прежде всего специализированными бактериями.

По сравнению с первыми клетками у бактерий значительно развилась информационная составляющая жизнедеятельности.

Основные составляющие генетических процессов были “выработаны”, вероятно, еще в доклеточный период (особенно фиксация кода нуклеотидов, взаимоотношения нуклеиновых кислот с белками и другое), они служат определяющими признаками жизни и имеются у всех организмов. Однако у бактерий эти составляющие заключены в клетке (какой бы примитивной она не была) — главном составляющем жизни. Кроме того, поведение бактерий уже имеет и внешнюю сигнальную основу (различные таксисы).

Бактерии находятся на грани живого и в смысле их размеров, которые могут быть предельно малыми: до 200 нм с толщиной в несколько нанометров (за пределами светового микроскопа). В тоже время существуют и многоклеточные бактерии (например, цианобактерии).

Для поддержания жизни необходимо обеспечить две основные функции: питание (структурное и энергетическое) и размножение. То и другое требует у бактерий существенного развития информационной составляющей жизненных процессов.

8. Энергетика и упорядоченность. Фундаментальные энергетические механизмы были освоены бактериями, вероятно, уже около трех миллиардов лет назад. Сутью характерной для жизни упорядоченности является информационная составляющая, которая позволяет в важные моменты жизнедеятельности организма делать “правильные” выборы. Эти выборы в критические моменты жизнедеятельности клетки направляются планом организма — геномом, а не случайным раскладом физико-химических условий, по крайней мере, если говорить о статистической картине жизни. Энергетическое обеспечение жизнедеятельности тесно связано с информационной составляющей жизненных процессов.

9. Термодинамика информативности. Стационарные состояния в клетке относятся к диссипативному типу, однако они существенно отличны от состояний, рассматриваемых синергетиками, вроде колец Бенара или “брюсселятора”. В синергетике отсутствует информационная составляющая, включающая в себя память, сигнальность, языковые структуры, основанные на геометрической форме элементов, спланированное управление, основанное на сигнальности.

Другая важная часть этой составляющей — квантованность. Процессы в живой клетке происходят на нескольких уровнях, из которых нижний уровень — уровень атомов и электронов. (Это значит, что описание жизненных процессов в клетке только на молекулярном уровне неполно, а уровень, скажем, кварков не нужен.) Наиболее элементарные жизненные механизмы не чисто статистические: в них ведется учет каждого электрона и протона, обрабатываются индивидуальные электроны, атомы, ионы и фотоны.

Примерами таких механизмов служат мембранные каналы, перенос электронов, молекулы АТФ, фотохимические центры зрительных рецепторов. Мембранные каналы могут быть пассивными (“конститутивными” — не требовать затраты энергии на прохождение через них ионов или радикалов) или активными (с переносчиками, потребляющими энергию, во всяком случае они избирательны). Цепи транспорта электронов представляют собой цепочки ферментов, встроенные в мембраны, которые буквально последовательно передают электроны один за другим, “поштучно”. Молекула АТФ (вездесущий переносчик энергии) является изобретением эволюции, которое благодаря расчетливой отмеренности кванта энергии в каждом элементарном процессе позволяет существенно уменьшать статистически обусловленные потери, передавая энергию с очень высокой эффективностью.

Клеточные процессы, направленные на построение структуры, как правило, идут против “естественного хода вещей”, т.е. энергетически невыгодны: на их протекание должна затрачиваться свободная энергия. Такие процессы должны идти в сопряжении с реакциями, в которых свободная энергия выделяется (и часть ее передается нужной реакции, а часть обращается в тепло).

Для осуществления нужной организму в данный момент цепи биохимических реакции требуется выполнение нескольких условий. Во-первых, нужно ферментативное ускорение реакций, т.е. одновременное появление в клетке набора соответствующих ферментов. Во-вторых, на каждом шаге этой цепи должна быть израсходована определенная порция энергии, достаточная и не чрезмерно большая: общая величина затрат энергии в клетке должна быть ограничена, и система должна избегать перегрева. В-третьих, трехмерная диффузия слишком медлительна для улавливания реагирующих молекул, которые к тому же должны быть повернуты определенным образом для попадания в активные центры ферментов. Поэтому основные реакции в клетке происходят на двумерных мембранах или даже на одномерных фибриллах, или с закреплением реагентов в сетях таких фибрилл (последнее — достижение эукариотов).

По-видимому, уже первичные организмы накапливали и потребляли энергию в форме молекул АТФ (или близкого типа). Во всяком случае эволюция механизмов потребления энергии и ее информационного обеспечения происходила в прокариотический период.

10. Прокариоты и эукариоты. Организация генов у прокариот (бактерий) и эукариот (организмов, у которых клетка имеет ядро) оказалась разной. Во-первых, как показали французские ученые Ф. Жакоб и Ж. Моно у прокариот гены собраны в группы, называемые *операми*. С оперона считается одна длинная молекула информационной РНК (мРНК), на которой затем синтезируется несколько белков, участвующих обычно в единой цепочке биохимических реакций. У эукариот каждый ген регулируется отдельно, на нем синтезируется своя молекула мРНК, которая кодирует только один белок. Это открывает более широкие возможности для процесса регуляции работы генов, так как теперь каждым геном можно управлять независимо от других. Во-вторых, у бактерий длина участка ДНК, с которого считывается мРНК, и длина самой молекулы мРНК одинаковы. У эукариот длина участка ДНК, с которого считывается мРНК, как правило, заметно больше, чем длина самой молекулы мРНК. Оказалось, что в таких генах эукариот на молекуле ДНК имеются участки, кодирующие

белки (их называют *экзонами*), разделенные вставками (их называют *интронами*), которые ничего не кодируют. В некоторых генах имеется до 50 интронов. Их суммарная длина может в 10 раз превышать длину экзонов.

Сначала с ДНК эукариот считывается длинная молекула РНК, в которой записаны и экзоны и интроны. Однако затем эта РНК видоизменяется: из нее вырезаются интроны и получается зрелая молекула мРНК. Этот процесс называется *сплайсингом*. Образно говоря, маленькие бактерии экономят место в своем геноме, не допуская бесполезных участков ДНК и имея общее начало для нескольких генов. В этом отношении вирусы, имеющие еще меньшие размеры, пошли дальше бактерий. У некоторых вирусов найдены гены, которые кодируют три разных белка. Какой белок получится, зависит от того, с какого нуклеотида начинается считывание. Эукариоты, имеющие более крупные клетки, чем бактерии, не столь экономят место.

У архебактерий — недавно открытой группы древнейших бактерий — тоже имеются интроны. Это позволяет предположить, что архебактерии были предками и других бактерий и эукариот. Но бактерии пошли по пути экономии генетического материала и утратили интроны, а эукариоты их сохранили. Возможно, хотя сами интроны обычно бесполезны, организация гена из кусочков, которые могут комбинироваться при перестройке ДНК, оказалась полезной.

7. ПРИЛОЖЕНИЕ К §3.

18. Регулон и оперон. Жизнь клетки (бактерии) между ее репликациями, т.е. вся ее собственная жизнь, состоит во включении и выключении групп генов (оперонов, регулонов), ответственных за определенные участки метаболической карты (т.е. совокупности всех биохимических реакций) организма, организацию транспорта молекул в клетку и т.п. Типично представлять себе такую картину: один метаболический путь (т.е. цепочка последовательно протекающих химических реакций, оканчивающаяся производством одного из сравнительно небольшого числа жизненно важных для клетки веществ) обеспечивается поступлением в клетку фиксированного набора ферментов, что, в свою очередь, обеспечивается экспрессией набора генов, точно соответствующих этим ферментам. Этот набор генов, часто называемый *регулоном*, совместно включается и выключается. Раньше здесь даже приводили тезис: “один ген — один белок-фермент”. Сейчас известно, что имеются гены, кодирующие производство РНК; регуляторные гены, цель которых посадить рибосому на образующийся соответствующий участок РНК для регуляции производства РНК (такой ген называется *лидерным*; об этом говорится ниже, в пункте 20); и т.д.

Регулон в случае его транскрибирования на одной молекуле РНК называется еще *опероном*, и далее мы говорим об этом последнем случае.

Обычно оперон имеет одну расположенную перед его первым геном регуляторную область, т.е. такой участок ДНК, который ответственен за включение и выключение этого оперона. Механизмы такого включения и выключения известны (они называются *механизмами* или *системами регуляции*); их немного и, возможно, что известны уже все системы регуляции в клетках бактерий.

19. Вторичная структура и спираль. Вторичная структура состоит из определенным образом расположенных *спиралей*, например таких, которые показаны на рис. 1. А спираль состоит из двух упорядоченных наборов (каждый из одинакового числа) *отрезков* нуклеотидов — левого и правого наборов. В каждом из этих наборов между соседними отрезками имеются какие-то *промежутки* — либо не имеющие парного промежутка в другом наборе (тогда этот промежуток называется *выпячиванием*), либо имеющие таковой (тогда эта пара промежутков называется *внутренней петлей*). Промежуток между самими наборами называется *внешней петлей*. Каждый *i*-ый отрезок, считая от начала левого набора должен быть комплементарным *i*-ому отрезку от конца правого набора.

Эти отрезки часто нумеруются именно в таком порядке, но иногда нумеруются и комплементарными парами, начиная от внешней петли. Эта вторая нумерация связана с представлением о том, что

В зависимости от этого один-ноль сигнала (двузначная логика) по разному пойдут процессы в биохимической карте организма. Таким образом, вторичная структура — это весьма сложный комбинаторный объект, и алгоритмическая работа с ним в рамках нескольких полных геномов родственных организмов является далеко не решенной математической задачей.

20. Аттенуаторная регуляция экспрессии генов. В этом пункте обсудим одну сравнительно новую систему регуляции, называемую *аттенуаторной регуляцией*. Ее пример показан на рис. 1, где представлена регуляторная область регулона *pheA*, ответственного за производство в клетке *фенилаланина* (одной из аминокислот).

Аттенуаторная регуляция имеет несколько вариантов, определяемых параметрами, роль которых играют число спиралей, участвующих во вторичной структуре, уровень, на котором происходит сама регуляция (транскрипции или трансляции), выбор индуктора, выбор вспомогательных молекул, участвующих в стабилизации конформаций, и т.д.

Суть этого механизма регуляции (в одном из его вариантов) состоит в “соревновании” рибосомы и РНК-полимеразы во взаимной скорости их прохождения по регуляторной области оперона.

Говоря более подробно, этот механизм таков. Слева направо в ДНК и соответственно в участке РНК, уже транскрибированном РНК-полимеразой, появляются три потенциальные спирали, перечисленные в пункте 19.

После паузной спирали (перед плечом 3, рис. 1) РНК-полимераза (которая до этого села на промоторе Р) делает паузу фиксированной длительности (не зависящей от других событий). Рибосома (которая позже села на область *Шайн-Дельгарно*, образовавшегося начального участка мРНК) быстро проходит *лидерный ген*, специально для этого приготовленный в этой регуляторной области (о таком гене говорилось в пункте 18). В составе лидерного гена находится *поле специфических* (для данного регулона) *кодонав*, в данном случае это несколько кодонав ttt и ttc, — на них рибосома ждет поступления из клетки соответствующей аминокислоты, в нашем случае она ждет фенилаланина. Таким образом, рибосома проходит лидерный ген быстро или медленно в зависимости от того много или мало фенилаланина в клетке. В *первом случае*, когда его много, рибосома проходит быстро и оказывается в точке стопа лидерного гена как раз в тот момент, когда двинувшаяся РНК-полимераза пройдет плечо 4; при этом рибосома стерически не дает создаться ближе к ней расположенной антитерминаторной спирали 2'-3, но не мешает созданию чуть дальше расположенной терминаторной спирали 3'-4. Эта вторая спираль создается, что является сигналом прекращения транскрипции: РНК-полимераза “соскакивает с ДНК” и не создается полноценная (синоним: *зрелая*) мРНК соответствующего оперона. (Спирали 2'-3 и 3'-4 незначительно отличаются от спиралей 2-3 и 3-4 соответственно.)

Во *втором случае*, когда фенилаланина мало, рибосома ждет его и задерживается на поле этих кодонав. И не успевает в точку стопа к моменту, когда РНК-полимераза пройдет плечо 4 (и даже цепочку нуклеотидов Т, которая идет за ним). Тогда из двух спиралей 2-3 и 3-4 складывается антитерминаторная спираль 2-3: она образовалась раньше спирали 3-4 и, как энергетически более выгодная, привела к расплетанию спирали 1-2 и по той же энергетической причине сама не расплелась, когда образовалась уже и спираль 3-4. Создание этой антитерминаторной спирали является сигналом для того, чтобы РНК-полимераза шла по оперону до конца. В результате образуется зрелая мРНК.

Здесь мы видим пример поразительной пространственно-временной синхронизации двух подготовленных процессов в регуляторной области оперона: движение рибосомы и движение РНК-полимеразы.

Закончим этот пункт кратким списком различных случаев аттенуаторных регуляций.

1. В классе Грам-отрицательных бактерий (например, для гамма-протеобактерий, в частности для *E.coli*) регуляция происходит на уровне элонгации (т.е. разворачивания процесса) транскрипции по типу *аттенуации транскрипции* с использованием трех-шпилечной вторичной структуры, как это было описано выше и показано на рис. 1.

Также регулируется биосинтез ароматических аминокислот (триптофана, фенилаланина, тирозина), гистидина, трианина и т.д., а также аминоксил-тРНК синтетаз. Регуляция генов транспорта аминокислот по такому механизму неизвестна. Медиатором регуляции является, как мы видели, рибосома, имеется лидерный ген, паузная спираль всегда расплетается.

2. В классе Грам-положительных бактерий (например, для группы *Lact. Lactis* и многих других) аттенуация транскрипции происходит с помощью вторичной *структуры Т-бокса*. Так регулируется биосинтез аминокислот (например, триптофана), аминоксил-тРНК синтетаз и многого другого, так регулируются гены транспорта. Медиатором служит тРНК в форме Г (не в форме клеверного листа). Удивительно, что такая тРНК двумя концами (пусть акцептор не занят) попадает в два определенных сайта; вероятно, здесь играют роль уже третичные конформации Т-бокса и тРНК. Имеются две конформации и из энергетических соображений одна из них (вида 1-2=тРНК=3-4) выгоднее, чем другая (вида 4'-5); лидерного гена нет, *спесифаер* (спираль аналогичная в этом случае паузной) всегда не расплетается.

Удивительно, что акцептор спаривается с Т-боксом водородной, а с аминокислотой 1-ковалентной связью (при участии фермента — аминоксил-тРНК синтетазы, своей для каждой аминокислоты и каждого вида тРНК). Это означает, например, что комплекс триптофан-тРНК работает в одном режиме, если триптофана много, и в другом, если триптофана мало.

Механизм взаимодействия состоит в том, что РНК-полимераза делает паузу после плеча 4'. Незагруженная тРНК "тычется" во все места и, когда ее много (т.е. триптофана мало), садится сразу на два сайта. Тогда образуется энергетически более выгодная конформация (терминирующая) 1-2=тРНК=3-4 (к тому же образовавшаяся чуть раньше) по сравнению с 1-2 — 4'-5 (антитерминирующей).

3. В классе Грам-положительных бактерий (например, для *B. Subt.*) аттенуация транскрипции осуществляется уже двух-шпилечной структурой. Так регулируется, например, биосинтез и транспорт триптофана. Медиатором служит ТРАП-белок, происходящий от специального гена *mtgV*. Происходит прямое действие триптофана через ТРАП на конформацию, а именно ТРАП стерически препятствует образованию более выгодной спирали 1-2.

Механизм состоит в следующем: если триптофана много, то он, соединяясь с белком ТРАП в комплекс, активизирует его. Затем комплекс садится на сайт, сильно перекрывающийся с плечом 1, что, конечно, не дает образоваться энергетически более выгодной спирали 1-2. Образуется спираль 2-3 и наступает терминация. Если плечо 1 не занято, то образуется спираль 1-2 и наступает антитерминация.

4. В классе Грам-положительных бактерий (например, для *Actinomycetes*, *Thermotoga*, *Thermus*, *Deinococcus* — на уровне *транскрипции*, а в классе Грам-отрицательных для многих протеобактерии, например для *Escherichia coli*, *Vibrio cholerae*, *Pseudomonas putida* и других, — на уровне *трансляции*) происходит регуляция с помощью RFN-структуры. Медиатором служит RFN-элемент, в этом случае лидерного гена нет.

Механизм состоит в прямом действии через молекулы FMN (которые садятся в петлях RFN-элемента, играя роль лиганда) на образование одной из двух конформаций, связанных с регуляторной областью: антитерминация или терминация в сочетании с цепочкой нуклеотидов Т (случай аттенуации транскрипции). Или связанных с Шайн-Дельгарно сайтом (это — последовательность AGGAGG, возможно, без первой или последней букв) и с посадкой на него рибосомы (случай аттенуации трансляции, так называемый, *ШД-секвестр*).

21. Эубактерии в противоположность архебактериям. Архебактерии были открыты в 1977 году. Считается, что архебактерии по сравнению с эубактериями (так сказать, классическими бактериями) примитивнее и ближе к общему предку прокариот и эукариот. Формы клеток архебактерий в целом сходны с таковыми у эубактерий. Среди них есть кокки, палочки, извитые клетки и виды, характеризующиеся слабым ветвлением. По тонкому строению клетки, выявляемому с помощью электронного микроскопа, архебактерии принципиально не отличаются от эубактерий и ближе к их грамполо-

жительной ветви. Прокариотная организация археобактерий проявляется в отсутствии у них ядра и характерных для эукариот органелл, окруженных мембраной. Хотя клетки археобактерий структурно относятся к прокариотному типу, многие макромолекулы, входящие в их состав (липиды, полисахариды, белки), уникальны и не найдены ни у эубактерий, ни у эукариот. С другой стороны, недавно у археобактерий описан известный только у эукариот процессинг: вырезание из первичного продукта транскрипции оперона определенных нуклеотидных участков, его укорачивание и в результате образование зрелых молекул РНК. Состав рибосомальных РНК археобактерий типично эубактериальный (5S, 16S и 23S рРНК), но их первичные структуры отличны от эубактериальных и эукариотных.

Существенные отличия выявлены у археобактерий и в строении генома, аппаратов репликации, транскрипции и трансляции.

Авторы выражают искреннюю признательность М.Б.Беркинблиту за многочисленные полезные советы.

СПИСОК ЛИТЕРАТУРЫ

1. Кузнецов Н.А. Информационные процессы в технических и живых системах. *Информационные процессы*. 2001, том 1, № 1, сс. 1-99.
2. Албертс Б. и др. *Молекулярная биология клетки*. М: Мир, 1994.
3. Любецкий В.А. Логические модели процессов информационного взаимодействия. *Труды международной конференции "Проблемы управления и моделирования в сложных системах"*, РАН, Самара, 1999, сс. 133–139.
4. Кузнецов Н.А., Любецкий В.А., Чернавский А.В. К вопросу о понятии информационного взаимодействия, 1: допсихическая биоинформатика. *Труды международной конференции "Проблемы управления и моделирования в сложных системах"*, РАН, Самара, 2000, сс. 8–20.
5. Кузнецов Н.А., Любецкий В.А., Чернавский А.В. К вопросу о понятии информационного взаимодействия, 2: доречевой интеллект. *Труды международной конференции "Проблемы управления и моделирования в сложных системах"*, РАН, Самара, 2001, сс. 25–42.
6. Кузнецов Н.А., Любецкий В.А., Чернавский А.В. К вопросу о понятии информационного взаимодействия, 3: речевой интеллект. *Труды международной конференции "Проблемы управления и моделирования в сложных системах"*, РАН, Самара, 2002, сс. 7–17.

(Подробный список литературы будет приведен в последующей публикации)