

ДИСКРЕТНОЕ МОДЕЛИРОВАНИЕ В БИОЛОГИИ РАЗВИТИЯ

В. З. Аладьев

Развитие организмов, как известно, является достаточно загадочным процессом. Как может из единственной клетки — оплодотворенного яйца — вырасти организм, состоящий из многих миллионов клеток, образующих чрезвычайно сложную саморегулируемую систему? Преклонение перед этим процессом еще более возрастает, если вспомнить, что он существенно автономен, что все клетки в организме генетически идентичны и что развитие строго контролируется изнутри. Говоря об автономности процесса, мы имеем в виду тот факт, что вся информация, необходимая для развития организма, содержится в исходной клетке; внешняя среда обеспечивает развитие энергией и материалами, а не информацией. Действительно, зигота определенного вида всегда превращается в организм этого же вида — каково бы ни было окружение. Рост осуществляется, главным образом, через непрерывный процесс самовоспроизведения клеток в организме, но дифференциация клеток в процессе роста более трудна для понимания, так как, по общему мнению биологов, все клетки содержат один и тот же набор генетических инструкций — новые клетки генотипично идентичны своим предшественникам. В связи с этим возникает вопрос: каким образом они становятся отличными друг от друга и развиваются в тщательно выработанные пространственные формы? Более того, весь процесс развития строго контролируется так, что различные системы организма развиваются в определенных пропорциях и во многих случаях организм способен преодолевать повреждения. Естественно, что процесс развития осуществляется на основе строгих механизмов контроля и адаптации. Пока мы не знаем лучшего подхода к выяснению всех вопросов, кроме как решить аналогичные проблемы для искусственных систем. Следует отметить, однако, что изучение феномена развития в организме привело некоторых исследователей [Driesch, 1929; Elsasser, 1958] к выводу, что организм нельзя рассматривать как машину.

С точки зрения кибернетики, общей теории систем и самой биологии весьма важно попытаться выяснить вопрос: может ли машина вообще развиваться подобно живым системам и, если может, то каким образом? Это важно знать по двум основным причинам: во-первых, если машина не может развиваться, то сохраняет

определенную силу аргумент о том, что живые системы обладают некоторым специфическим феноменом. В этом случае аргумент кибернетики о том, что живые и неживые системы могут быть определены в терминах одних и тех же принципов, был бы поставлен под сомнение. Во-вторых, при положительном ответе, т. е. если были бы поняты принципы развития неживых систем и проведена удовлетворительная аналогия с живыми системами, то наряду с важными революционизирующими применениями в технологии многих производственных процессов, мы смогли бы получить удовлетворительный аппарат исследования живых развивающихся систем. Рассмотрим основные попытки в этом направлении.

Развитие организма обычно состоит из роста и дифференциации. Рост, как известно, означает простое увеличение размера организма главным образом за счет самовоспроизведения клеток. Дифференциация является более сложным процессом, и целесообразно выделять по крайней мере два ее типа: пространственный и фенотипический, который Аптер [1970] называет функциональным. Так, в растущей ткани можно выделить изменение формы и конфигурации межклеточной связи (пространственная дифференциация) наряду с увеличением дифференциации отдельных типов клеток (фенотипическая). Следует отметить, что для пространственной дифференциации в биологии существует устоявшийся термин «морфогенез», однако для целей моделирования более естественным представляется именно первый термин. Конечно, фенотипическая дифференциация также имеет место в пространственной дифференциации, но для простоты мы будем рассматривать их отдельно. Развивающийся организм характеризуется не только возможностью достичь сложной пространственной и фенотипической дифференциации, но он в большей или меньшей мере обладает способностью к регуляции и регенерации. Под регуляцией мы понимаем свойство организма развиваться в нормальную особь, даже если он подвержен изменениям в процессе развития (например, при удалении или перестройке клеток), тогда как под регенерацией понимается свойство организма восстанавливать любое нарушение, которое организм получил в момент своего полного развития.

Несмотря на важность понимания биологического развития, включающего пространственную и фенотипическую дифференциацию, регуляцию, регенерацию, а также феномен самовоспроизведения, попытки достичь успеха в моделировании этого процесса можно отнести к первому этапу модельного периода, который характеризуется моделированием отдельных феноменов процесса развития, причем для моделирования использовалась самая разнообразная техника. Общим для всех этих моделей являлся сам принцип исследования: формализация изучаемого феномена → → построение конкретной модели → сравнительный анализ функционирования модели и реального биологического феномена. Основную роль первого этапа моделирования можно охарактери-

звать тем, что целому ряду сложных процессов развития была дана удовлетворительная формализация, которая корректировалась на основе анализа многочисленных формальных моделей [Аладьев, 1972; Aladyev, 1974—1976, 1978, 1980]. Анализ ряда моделей позволил по-новому взглянуть на некоторые регуляторные механизмы развития. Однако мы имели ряд не связанных общей теоретической базой моделей. Естественно, что такое положение не способствовало выработке единого аппарата моделирования биологии развития. Однако в рамках уже первого этапа зародились две техники моделирования ряда феноменов развития: клеточные автоматы и развивающиеся параллельные грамматики Линденмайера. Клеточные автоматы, известные затем как однородные структуры (ОС), были использованы фон Нейманом [1971] для изучения проблемы самовоспроизведения, а параллельные развивающиеся грамматики были впервые введены Линденмайером для моделирования морфогенеза [Lindenmayer, 1975, 1978; Herman, Rosenberg, 1975; Rosenberg, Salomaa, 1974; Rosenberg et al., 1977] и впоследствии получили название *L*-систем. ОС- и *L*-системы в настоящее время представляют собой наиболее общий и популярный аппарат кибернетического моделирования развития [Herman, Rosenberg, 1975; Aladyev, 1980].

Определим неформально понятие ОС. В каждую целочисленную точку *d*-мерного Евклидова пространства (Z^d) помещена копия одного и того же конечного автомата. Каждый из них связан с конечным числом соседних автоматов согласно некоторому индексу соседства (*X*), который одинаков для всех автоматов пространства. В каждый целочисленный момент времени *T* автомат изменяет свое состояние из конечного множества $A = \{0, 1, 2, \dots, a-1\}$ в зависимости от конфигурации состояний его самого и всех соседних автоматов в предыдущий момент $T - 1$. Изменения состояний автомата определяются локальной функцией перехода σ . Одновременное применение функции σ ко всем автоматам пространства определяет глобальную функцию перехода τ , которая преобразует одну конфигурацию пространства Z^d в другую. Среди всех возможных состояний автомата пространства выделяется так называемое состояние покоя (q_0), суть которого состоит в том, что автомат в состоянии q_0 не изменяет своего состояния в следующий момент, если все его соседи находились в состоянии покоя. Это состояние введено, чтобы наложить ограничение на скорость передачи информации в ОС. Таким образом, ОС есть упорядоченная пятерка $OS = (Z^d, A, X, \sigma, q_0)$; это так называемое классическое понятие ОС. В настоящее время математическая теория ОС представляет собой хорошо развитый аппарат исследования многих дискретных процессов [Aladyev, 1980; Аладьев, 1982].

Аппарат ОС позволяет формальными средствами исследовать на клеточном уровне такие феномены развития, как рост, самовоспроизведение, дифференциация, регуляция и регенерация. В настоящее время ОС позволили реализовать целый ряд моделей

развития, получающих весьма интересные биологические интерпретации [Aladyev, 1980; Аладьев, 1981]. Наряду с этими проблемами ОС позволяют удовлетворительно исследовать такие вопросы развития, как сложность развивающихся систем, процессы, управляющие ростом, регуляцией и регенерацией, устойчивость процессов развития, необходимые и достаточные условия регуляции и регенерации и т. д. [Аладьев, 1981]. Однако наряду с этим ОС порождают трудности при моделировании в них ряда биологических феноменов. Основные трудности связаны с большой чувствительностью ОС к размерности моделей и с ограничениями на возможность деления клеток внутри моделируемого развивающегося организма.

Учитывая трудности моделирования в ОС ряда биологических феноменов, Линденмайер [Lindenmayer, 1975] ввел свои системы, известные сейчас под названием «систем Линденмайера» (*L*-систем). В рамках *L*-систем для моделирования морфогенеза и растущих структур Линденмайер предложил ветвящиеся алгоритмы [Lindenmayer, 1978], а рядом авторов (Кулик, Линденмайер, Нагль и др. — ссылки см. в обзоре: Lindenmayer, Culik, 1981) для моделирования развития и роста были введены графические порождающие системы. На базе *L*-систем был реализован целый ряд растущих алгоритмов, обзор которых может быть найден в прекрасной работе Линденмайера [Lindenmayer, 1975]. Г. Люк и Дж. Люк [H. Lück, J. Lück, 1978] для объяснения роста тканей также использовали модель на основе *L*-систем. В последнее время на основе *L*-систем появляется все большее число моделей как собственно роста, так и роста в составе общего феномена развития [Nishio, 1978]. Поэтому уместно ввести понятие *L*-систем.

Пусть V есть конечный алфавит множества состояний клетки, а V^* обозначает множество всех одномерных слов из элементов V . V^* включает и пустое слово, обозначаемое через Λ . Тогда *L*-система есть упорядоченная четверка $\langle V, b, \delta, bS \rangle$, где $b \in V$ — постоянное воздействие внешней среды на слово (клетку); σ — функция, дающая по состояниям любых трех прилегающих клеток непустое конечное множество слов в алфавите $V \setminus \{b\}$, на которые заменяется внутренняя клетка из трех прилегающих; bS — начальная клетка, из которой начинается процесс развития. Таким образом, рассматривается система параллельных правил δ , которые порождают из одного начального однобуквенного слова (зиготы) множество одномерных слов (имитирующих развитие организма) в некотором конечном алфавите V . Правила порождения слов разрешают вставки любого количества букв в любые места слова. При этом допускаются и недетерминированные переходы.

Таким образом, *L*-системы существенно расширяют одномерные ОС в смысле множества порождаемых слов. С точки зрения биологической адекватности они получают вполне удовлетворительные интерпретации. *L*-системы уже хорошо зарекомендовали себя при описании целого ряда биологических процессов и в на-

стоящее время, по всей вероятности, представляют собой наиболее разработанный в математическом плане и адекватный в биологическом плане аппарат для моделирования биологии развития. В отношении самого аппарата L -системы более абстрактны, чем ОС, хотя бы потому, что они не привязаны жестко к системе координат и, по сути дела, являются одним из типов параллельных формальных грамматик, которые сейчас интенсивно изучаются [Rosenberg et al., 1977]. Следует отметить, что и ОС могут рассматриваться как некоторый тип параллельных грамматик (τ_n -грамматики) [Аладьев, 1981], которые являются собственным подклассом класса всех L -грамматик. Ниже мы детальнее проанализируем ОС и L -системы на предмет их возможностей биологического моделирования.

Основой биологических интерпретаций ОС стали следующие предположения.

1. В качестве биологической единицы больше всего подходит клетка, которая есть некоторый клеточный автомат, и все, что нам нужно знать о ней, — это зависимость ее выхода от входа и ее состояния.

2. Все клетки в организме имеют одинаковый генотип, т. е. один и тот же набор инструкций ее работы.

3. Развитие системы клеток существенно зависит от обмена информацией между этими клетками.

4. Организм сам регулирует наиболее важные стороны своего развития. Иными словами, развитие управляется изнутри, а не извне.

Конечно, каждое из этих четырех предположений является упрощением реального положения вещей, но, когда модели, основанные на них, помогут добиться какой-то ясности, в них можно будет включить новые предположения, чтобы приблизить их больше к реальности. Так, при моделировании регуляции, регенерации и дифференциации в ОС мы дополнительно к указанным предположениям использовали и принцип Сэджер об образовании форм в соответствии с системой инструкций. Это представление об инструктивной системе наиболее привлекательно именно тем, что оно может служить развитием того пути, по которому обычно мыслится приложение теории информации в биологии развития.

Нетрудно убедиться, что поведение конечных ОС может быть описано на языке логических сетей. А так как Сугита [Sugita, 1963] доказал возможность выражения моделей Жакоба—Моно на языке логических сетей, то модели развития в ОС могут получить генетическую интерпретацию на языке моделей Жакоба—Моно. Решение этой задачи и результаты Сугита докажут тогда эквивалентность логических сетей и сетей Жакоба—Моно, из чего будет следовать и принципиальная возможность интерпретации моделей, реализованных в ОС, моделями Жакоба—Моно. В этом же направлении весьма интересно применить и работу Сталя и Гозна [Stahl, Goheen, 1963].

Моделирование в ОС позволяет рассматривать процессы раз-

вития с точки зрения иерархических структур [Аладьев, 1981]. Таким образом, модели развития, реализованные в ОС, могут исследоваться средствами систем, что обеспечивается свойствами ОС. Такой подход позволяет получать качественно новые результаты моделирования процессов развития.

Для моделирования процесса образования осевой структуры, который является составной частью общего процесса развития, мы использовали несколько типов ОС [Аладьев, 1972; Aladyev, 1980]. Но при обсуждении моделирования общей проблемы развития к четырем упомянутым выше предположениям необходимо добавить пятое: развитие каждого организма осуществляется самовоспроизведением составляющих его клеток. Именно этим пяти основным предположениям очень хорошо отвечают ОС. В этом случае клеткам развивающегося организма отвечают автоматы ОС. На самом деле реальная клетка по своей структуре намного сложнее наших автоматов, но для изучения проблемы на клеточном уровне мы должны как-то упростить проблемы, связанные со структурной сложностью клетки, а также с работой ее составных частей. Этого мы и добиваемся, рассматривая клетку как черный ящик — клеточный автомат.

Следовательно, рассматривая межклеточные взаимодействия, переходим на более высокий уровень организации, чем тот, который исследовали Сталь и Гоэн [Stahl, Goheen, 1963], моделируя системы ферментов на основе машин Тьюринга. Таким образом, за единицу мы берем клетку. И в смысле ее поведения ограничиваемся только зависимостью ее выхода от входа и ее состояния в предыдущий момент времени. В принципе, ОС вполне допускают и структурный уровень моделирования, когда при реализации в ОС моделей развития исследуется и внутренняя организация единичного автомата. В ряде моделей развития, упоминавшихся выше, уже использованы внутренние структуры единичного автомата ОС. Таким образом, ОС допускают моделирование процессов и на более низких уровнях, чем отдельная клетка.

Время в ОС полагается дискретным, тогда как на самом деле оно непрерывно. Но для целей кибернетического моделирования процессов развития это пока несущественно. Каждая клетка имеет один и тот же генотип, от которого, по-видимому, и зависит вид получившегося организма. Поэтому, вероятно, наиболее удобным путем создания эффективных моделей развития является моделирование генотипа (программы работы клетки). Действительно, до сих пор основная часть работы по дифференцировке клеток велась на уровне взаимодействия между тканями, тогда как крайне существенно распространить наше понимание на природу процессов, происходящих на клеточном уровне. Таким подходом мы выбрасываем из рассмотрения одноклеточные организмы, также испытывающие развитие, которое в основном есть результат внутриклеточной деятельности. Однако на нашем этапе моделирования мы пока вынуждены с этим мириться. Выше мы предположили, что все клетки имеют один и тот же генотип, т. е. каждая клетка в организме,

начиная с зиготы, приступает к работе с одним и тем же набором генетических инструкций. Само понятие ОС включает в себя именно этот аспект. Мы остановимся на этом моменте несколько подробнее.

Имеются данные о том, что в ряде организмов у различных клеток может быть разный генотип. Однако из самого определения ОС не следует, что все клетки моделируемого на таких структурах развивающегося организма идентичны. Действительно, при наличии одной и той же программы работы автоматы ОС в результате дифференцировки (изменения внутренних состояний) в разных областях однородного пространства имеют различные внутренние состояния (фенотипы) и в связи с этим по-разному реагируют на одинаковые входные сигналы. Таким образом, дифференцировке фенотипов клеток в ОС соответствует дифференцировка внутренних состояний автоматов. В связи с этим на каждом этапе развития ОС или какая-то ее часть, дифференцируясь, приобретает в общем случае новые свойства. Вообще термин «дифференцировка» имеет ряд толкований. Довольно интересное толкование этого термина дано Аптером [1970]. С точки зрения типов дифференцировки наибольший интерес в первую очередь представляет регионализация, или, как считает Аптер [1970], основной вопрос состоит скорее в выяснении того, как структура возникает сначала на уровне организации (как происходит «саморазметка» клеток), чем в установлении природы физического механизма, осуществляющего этот план, хотя этот механизм и может даже приводить к увеличению сложности. ОС позволяют получить конструктивные ответы на подобные вопросы.

Выше мы предполагали, что развитие организма существенно зависит от обмена информацией между клетками. В настоящее время этот факт является общепризнанным и имеет целый ряд подтверждений в свою пользу [Аптер, 1970; Aladyev, 1974]. В случае ОС таким обменом информацией служит передача одним автоматом другому (другим) символов из входного (выходного) алфавитов или сообщение о своем состоянии. На наш взгляд, нет необходимости в каком-либо ограничении для передачи информации от автомата к автомату в ОС ввиду сложности межклеточных взаимодействий (особенно химических) в реальных организмах.

Так как организм вырастает из единственной клетки (зиготы) путем самовоспроизведения этой клетки, то автомат ОС должен обладать способностью в нужный момент породить свою копию. Это достигается тем, что какой-нибудь соседний нефункционирующий автомат ОС переводится в некоторое ненулевое внутреннее состояние, становясь таким образом функционирующим и составной частью уже более сложного организма. Более того, предполагается, что как только нефункционирующий автомат станет функционирующим, то он автоматически получит весь генотип (программу работы) исходного автомата, из которого развивается организм в ОС.

На этом мы заканчиваем обсуждение свойств ОС с точки зрения упрощенных биологических предпосылок, лежащих в основе развития, и переходим к рассмотрению классов задач, относящихся к тому или иному процессу развития, которые можно моделировать и исследовать в ОС.

Первый класс задач включает в себе вопрос о том, каким образом осуществляются регуляция, регенерация и дифференциация в организме. Построенные для этой цели модели в ОС позволили прояснить ряд вопросов и сформулировать интересные проблемы для дальнейших исследований. Вопросы, относящиеся к этому классу задач развития, могут привести к лучшему пониманию проблемы образования пространственной структуры вообще. С другой стороны, в процессе решения этих вопросов понятие ОС было расширено и сделано более приемлемым для биологического моделирования.

Второй класс задач относится к исследованию процесса роста, который сам по себе не составляет проблемы с абстрактной точки зрения, так как обеспечивается самовоспроизведением клеток, на котором основано само развитие. Однако эта проблема включает вопрос, как организм может ограничивать свои размеры в процессе самовоспроизведения клеток, если это должно полностью обеспечиваться генотипом самой клетки. Действительно, интерес представляет такой рост, когда может происходить пространственная дифференцировка в процессе непрерывного самовоспроизведения исходного набора инструкций без влияния внешнего воздействия. Существенно также изучение процессов роста, которые приводят к получению из ограниченного числа инструкций организмов, состоящих из большого числа клеток. В этом плане перспективно исследование устойчивости процессов роста и их контролируемости по отношению к различного рода нарушениям, так как это может представить определенный интерес и для такой области, как онкология. Некоторые из перечисленных вопросов исследовались с помощью моделей роста, обсуждавшихся нами в предыдущем разделе.

Третий класс задач — это самовоспроизведение организмов. Существующие до сих пор модели самовоспроизведения характеризуются тем, что один организм строит свою копию. Однако с точки зрения развития наибольший интерес представляет именно вопрос: как может один автомат ОС, начав процесс самовоспроизведения, породить довольно сложный пространственно-дифференцированный нетривиальный организм, способный к самовоспроизведению и в определенной мере к регенерации? В связи с этим возникает вопрос: как процесс образования довольно сложного пространственного организма оберегает себя от ошибок и каким набором инструкций (алгоритмом) исходного автомата ОС этого можно добиться? Таким образом, круг задач третьего класса предполагает подход к самовоспроизведению на уровне клеток а не организмов.

Четвертый класс задач можно охарактеризовать как проблему сложности в биологии развития. Здесь можно сформулировать ряд весьма интересных вопросов о сложности автомата (клетки), из которого вырастает сложный многоклеточный организм, о сложности пространственной дифференцировки, об изменении сложности в процессе развития организма, да и о самом понятии сложности. Некоторые из этих вопросов исследовались в рамках моделей роста, но до полного понимания проблемы еще далеко.

Каждая из перечисленных задач четырех описанных классов является составной частью единого процесса развития, однако в настоящее время необходимо думать о том, как разложить развитие на составляющие и проанализировать их. В то же время мы не должны думать, что какая-то простая модель раскроет все свойства данного процесса развития. Поэтому каждый из классов задач предполагает целый ряд других возможных направлений для моделирования, и сам перечень классов может впоследствии быть расширен.

На начальном этапе кибернетического моделирования, когда вызывала сомнение сама принципиальная возможность, подобный подход с использованием ОС был, пожалуй, единственно возможным как в плане существования пригодного аппарата, так и в плане готовности самой биологии. Более того, сами ОС зародились именно как средство такого моделирования. Однако, моделируя на ОС процессы развития, мы в значительной мере игнорировали основу этих процессов — размножение клеток. Действительно, деление клеток в таких моделях могло осуществляться только на границах организма, а внутренние его клетки принципиально не имели такой возможности, так как ОС всегда жестко привязана к системе координат в пространстве E^n и вклинивание между клетками новой встречает непреодолимые трудности. Это в свою очередь не может не сказаться на качестве моделирования процессов развития в ОС. Поэтому уже сейчас можно отметить, что ОС (несмотря на то что они допускают ряд интересных обобщений, существенно расширяющих их возможности для моделирования биологии развития [Aladyev, 1980]) не смогут выполнить роль того универсального аппарата моделирования, с помощью которого можно было бы успешно исследовать биологию развития в целом. Подобное заключение можно сделать и относительно аппарата L -систем.

С целью получения более приемлемого аппарата моделирования Ле Хой [Le Hoi, 1978] ввел параллельные обменные системы, которые наследуют основные черты обеих систем (ОС- и L -систем), хотя и они весьма чувствительны к размерности пространства и имеют плохо формализуемые элементы (движение модулей в пространстве). Таким образом, вряд ли на основе упомянутых формальных аппаратов может получить развитие качественно новый аппарат, как это было в случае появления L -систем на базе одномерных ОС.

На основе анализа основных недостатков существующих аппаратов моделирования мы попытались выявить возможные пути разработки более адекватного аппарата моделирования процессов развития [Аладьев, 1981].

Это, во-первых, разработка специальных многомерных параллельных грамматик и алгоритмов, допускающих максимальное параллельное выполнение операций и вставку в любое место слова любых конечных подслов; во-вторых, использование идей и понятий графо-топологического аппарата и, в-третьих, разработка совершенно нового аппарата моделирования, наилучшим образом приспособленного для биологии развития.

На что же реально в настоящее время имеет смысл ориентироваться? Прежде всего о первой возможности. Разработка алгоритмов и грамматик, оперирующих с многомерными словами, представляет действительно большой интерес. Однако здесь опять возникают проблемы, связанные с размерностью. Наш поиск удовлетворяющих таким свойствам многомерных грамматик и алгоритмов, пока ни к чему не привел. Однако работа в этом направлении принесла свои положительные результаты для моделирования биологии развития [Аладьев, 1981].

Тем более не имеет смысла рассчитывать на создание в ближайшее время совершенно нового аппарата, наилучшим образом приспособленного к нуждам биологии развития. Как показал опыт, такой фундаментальный аппарат не создается быстро, да и притом на почти пустом месте. Остается возможность использования идей графо-топологического аппарата. Наиболее пригодным в настоящее время представляется именно дискретный графо-топологический подход, который действительно учитывает все указанные трудности предыдущих систем, хотя и является еще более абстрактным и менее наглядным. Более того, уже имеются некоторые модели развития, использующие этот подход. Среди них следует отметить интересные модели Аптера [Apter, 1973] и Линденмайера [Lindenmayer, 1978].

В заключение хотелось бы подчеркнуть, что мы разделяем мнение Уоддингтона, что основная теория должна быть сходна с топологией n -мерного пространства. Поэтому уже сейчас мы должны очертить наиболее существенные на клеточном уровне феномены многоклеточных развивающихся организмов, искать наиболее адекватный им математический аппарат и при описании процессов развития не ограничиваться обычными понятиями кибернетики и математики!