

**Математические
модели в биологии
и медицине**

Обсуждаются значения понятия универсальности при управлении внутриклеточными процессами жизнедеятельности. Проводятся параллели между внутриклеточными процессами, координируемыми белками и вычислениями, а сама клетка сравнивается с универсальным компьютером.

© Б.А. Белецкий, А.М. Гупал,
2015

УДК 577.1

Б.А. БЕЛЕЦКИЙ, А.М. ГУПАЛ

**УНИВЕРСАЛЬНОСТЬ
В МОДЕЛИРОВАНИИ
ВНУТРИКЛЕТОЧНЫХ ПРОЦЕССОВ**

Введение. Идея универсальности вычислений сыграла решающую роль в истории развития вычислительной техники. Она была разработана Тьюрингом в 1936 году в результате исследования проблемы разрешимости в виде так называемой универсальной машины Тьюринга [1]. Универсальная машина Тьюринга принимает на вход упорядоченную пару натуральных чисел, кодирующих соответственно входные данные и конкретную машину Тьюринга, работу которой универсальная машина Тьюринга имитирует на предоставленных входных данных. Таким образом, программное обеспечение было отделено от вычислительного устройства, что значительно ускорило развитие вычислительной техники.

Идея универсальности вычислений, предложенная Тьюрингом, была с энтузиазмом воспринята фон-Нейманом, и легла в основу архитектуры его универсального компьютера, которая наследуется сегодня большинством вычислительных устройств на Земле. Архитектура фон-Неймана позволяет перепрограммировать одно универсальное вычислительное устройство для решения широкого круга практических задач. Универсальный компьютер фон-Неймана состоит из:

- ПЗУ – постоянно-запоминающее устройство, позволяющее хранить информацию длительное время, независимо от компьютера;
- ОЗУ – оперативное запоминающее устройство, используемое для хранения промежуточных данных в процессе вычислений;

- устройства ввода/вывода – различные устройства, позволяющие обмениваться информацией с вычислительным устройством во время проведения вычислений;
- процессор – выполняет вычисления и координирует работу остальных компонент системы.

На рис. 1 показаны составные элементы архитектуры фон-Неймана и способ их взаимодействия. Вычисление в универсальном компьютере заключается в упорядоченном изменении состояний ячеек оперативной памяти, в результате чего ОЗУ компьютера переводится из некоторой начальной конфигурации в терминальную конфигурацию, сигнализирующую о завершении вычислений.

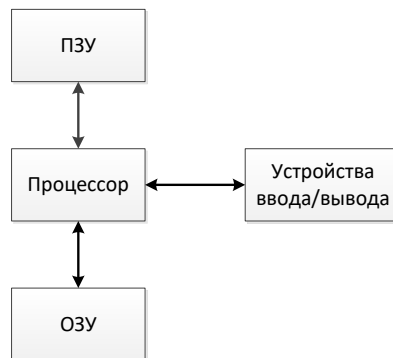


РИС. 1. Универсальный компьютер фон-Неймана

Жизненный цикл вычисления в универсальном компьютере состоит из следующих шагов:

- 1) инициализация – на этом этапе вычисления, хранящиеся в закодированном виде в памяти ПЗУ интерпретируются универсальным компьютером в виде в начальной конфигурации ячеек оперативной памяти;
- 2) выполнение – на этом этапе происходит последовательное изменение состояний ячеек оперативной памяти, в результате чего ОЗУ переводится из начальной конфигурации в терминальную;
- 3) завершение – освобождение ресурсов универсального компьютера, задействованных в процессе вычисления. Такими ресурсами являются ячейки ОЗУ, устройства ввода/вывода и области памяти ПЗУ.

Идея универсальности, лежащая в основе архитектуры фон-Неймана привела к революции в вычислительной технике, значительно упростив проектирование, реализацию, поддержку и выполнение вычислений. Поскольку оперировать текстом компьютерных программ значительно проще, чем деталями узкоспециализированных вычислительных устройств.

Универсальность в живой природе. Серия фундаментальных открытий в области молекулярной биологии, произведенных в прошлом веке, продемонстрировала важность понятия информации для процессов жизнедеятельности

организмов. Еще до открытия структуры молекулы ДНК и ее функции в клетке Шредингер предположил [2], что в хромосомах содержится описание организма и инструменты для его построения, благодаря чему сложный организм развивается из единственной клетки, удваивая и передавая хромосомы по наследству при каждом делении.

Фон-Нейман, изучавший самовоспроизводящиеся клеточные автоматы, уточнил идею Шредингера, предположив, что в ДНК находится описание организма, а также и описание инструментов его построения. Идеи фон-Неймана полностью подтвердились после открытия структуры молекулы ДНК и механизма экспрессии генов.

Выяснилось, что всеми живыми организмами на Земле используется один и тот же универсальный язык программирования внутриклеточных процессов, называемый генетическим кодом. Согласно генетическому коду инструкции, закодированные в молекуле ДНК, превращаются в активные внутриклеточные процессы.

На ранних этапах эволюции жизни на Земле произошла унификация внутриклеточных процессов жизнедеятельности, ограничив их многообразие достаточно богатым множеством процессов, которые координируются белками, построенными из 20 стандартных аминокислот. Таким образом, любому внутриклеточному процессу можно поставить в соответствие множество белков, участвующих в его координации. Структура любого белка, координирующего процессы жизнедеятельности клетки, закодирована в ДНК клетки (за исключением, возможно, тех белков, которые попали в клетку извне).

Внутриклеточные процессы жизнедеятельности можно сравнить с вычислительными процессами в компьютере, при этом роль ячеек оперативной памяти играют органические молекулы, находящиеся в цитоплазме живой клетки. Тогда саму живую клетку можно сравнить с универсальным компьютером, а ее важнейшим органеллам можно поставить в соответствие элементы архитектуры фон-Неймана. Действительно, важнейшими функциональными элементами живой клетки являются.

- ДНК – соответствует ПЗУ универсального компьютера, позволяя хранить в закодированном виде описание внутриклеточных процессов. Аналогично ПЗУ, где информация сохраняется даже после выключения компьютера, характерной особенностью ДНК является способность хранить генетическую информацию, вне зависимости от процессов жизнедеятельности клетки. Это способность, проявляется, например, при передаче генетической информации между клетками с помощью вирусов, которые не имеют собственного метаболизма и не являются живыми существами.

- Цитоплазма – ограниченная мембранной областью, заполненная растворенными в воде органическими молекулами. Клеточную цитоплазму можно сравнить с ОЗУ компьютера, в котором протекают вычислительные процессы. При внутриклеточных процессах вместо упорядоченного изменения конфигураций ячеек ОЗУ происходит упорядоченное изменение конфигураций органических молекул в цитоплазме.

- Белки – высокоактивные органические молекулы, которые подобно процессору компьютера, являются движущей силой внутриклеточных процессов. Именно под действием белковой активности происходят изменения конфигураций органических молекул в цитоплазме.

- Мембрана – защитный слой, обладающий избирательной пропускной способностью, и отделяющий внутриклеточное пространство от внешней среды. В мембране содержится множество белковых каналов, осуществляющих транспорт молекул сквозь мембрану; в этом смысле мембрану можно сравнить с устройствами ввода/вывода в универсальном компьютере.

На рис. 2 показаны важнейшие структурные элементы живой клетки, а также способы взаимодействия между ними. Жизненный цикл внутриклеточного процесса, подобно жизненному циклу цифрового вычисления на универсальном компьютере, можно разбить на 3 этапа.

1. Инициализация – на этом этапе РНК-код внутриклеточного процесса копируется с молекулы ДНК и интерпретируется в виде белковых структур в клеточной цитоплазме.

2. Выполнение – на этом этапе под действием высокоактивных белковых молекул происходит упорядоченное в пространстве и времени изменение конфигураций молекул в цитоплазме. В результате цитоплазма переводится из некоторого начального состояния в терминальное состояние.

3. Завершение – на этом этапе освобождаются ресурсы, выделенные для выполнения внутриклеточного процесса. Такими ресурсами в живой клетке могут быть отдельные стандартные аминокислотные остатки, из которых строятся белки, мембранные каналы или регуляторные участки генов.

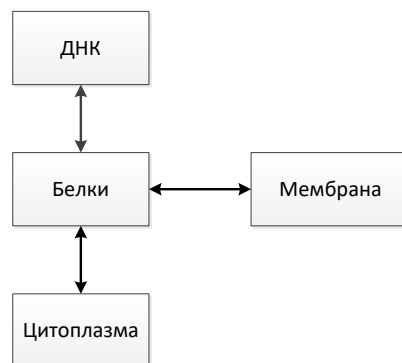


РИС. 2. Схема устройства живой клетки

Идея универсальности лежит в основе всех процессов жизнедеятельности, поскольку все внутриклеточные процессы запускаются по соответствующим ДНК-кодам. Механизм запуска внутриклеточных процессов называется механизмом экспрессии генов. Этот механизм присутствует у всех живыми организмами на Земле, и является универсальным в том смысле, что потенциально позволяет запускать любой внутриклеточный процесс, описанный в виде белков,

состоящих из стандартных аминокислот 20 типов (не ограничиваясь только теми процессами, которые закодированы в ДНК конкретного организма).

Универсальность – это одна из фундаментальных причин успешности и разнообразия форм жизни на Земле. ДНК-коды внутриклеточных процессов, присущих одному организму могут запускаться клетками других организмов, что позволяет живым организмам свободно обмениваться генетической информацией и образовывать новые виды.

Другой важной причиной использования универсальных процессов в клетке является то, что манипулировать внутриклеточными процессами значительно проще на уровне соответствующих ДНК-кодов, чем на уровне белковых молекул. Это связано с высокой активностью белковых молекул, а также с относительной инертностью молекул ДНК, закрученной в двойную спираль. Очевидно, ввиду этой особенности в процессе эволюции появилось множество механизмов направленной манипуляции молекулами ДНК, к которым относятся: поиск специфических последовательностей ДНК, копирование информации с ДНК, манипуляция (вырезание, вставка, замена) отдельными нуклеотидами ДНК, а также манипуляция (дупликация, удаление, вставка, транспозиция, изменение ориентации) целыми участками ДНК.

Моделирование внутриклеточных процессов с помощью частиц. В работе [3] предлагается моделировать внутриклеточные процессы и координирующие их белки с помощью частиц, обладающих двумя характеристиками: потенциалом и активностью. Потенциалы позволяют влиять на вероятности появления различных конфигураций частиц, а активность частиц определяет возможные способы преобразования этих конфигураций частиц.

Рассмотрим конечное множество B типов базовых частиц и конечное множество Γ двухместных связей вида $\gamma[\cdot, \cdot]$. Алфавит возможных типов частиц определяется индуктивно: 1) любой элемент из B является элементом A ; 2) любая n -местная связка вида $\gamma[\alpha_1, \alpha_2, \dots, \alpha_n]$, где $\gamma \in \Gamma$, $\alpha_1, \alpha_2, \dots, \alpha_n \in A$ является элементом A ; 3) ничто другое элементом A не является.

Частицы, определенные с помощью связей из Γ будем называть составными. Любую составную частицу $a \in A$ можно представить в виде дерева. Например, если $a = \gamma_1[\alpha_1, \gamma_2[\alpha_2, \alpha_3]]$ и $\alpha_1, \alpha_2, \alpha_3$ – некоторые базовые частицы, а γ_1, γ_2 – связки, то a соответствует дереву, показанное на рис. 3.

Вектор $\omega = (\omega_1, \dots, \omega_{|K|})$, $k \in K$, компоненты которого проиндексированы элементами конечного множества K и принимают значения из алфавита A , $\omega_k \in A$, будем называть конфигурацией системы. Множество всех конфигураций обозначим Ω . Структура конфигурации задается с помощью бинарного отношения соседства на множестве индексов K

$$N(k_1, k_2) = \begin{cases} 1, & \text{если } k_1 \text{ и } k_2 \text{ соседи;} \\ 0, & \text{иначе.} \end{cases}$$

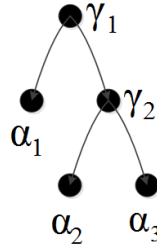


РИС. 3. Составные частицы, представленные в виде дерева

Рассмотрим множество U функций вида $u : \Omega \times K \mapsto \mathbf{R}$, которые задают локальные энергии конфигураций частиц. С каждой частицей $\alpha \in A$ связан неотрицательный вектор зарядов $\varepsilon(\alpha) = (\varepsilon_1, \dots, \varepsilon_{|U|}) \in \mathbf{R}^{|U|}$, а с каждой n -местной связкой $\gamma \in \Gamma$ связана функция $\varepsilon_\gamma : \mathbf{R}^{n|U|} \mapsto \mathbf{R}$ агрегации n векторов зарядов частиц, находящихся в связке, что позволяет вычислить вектор зарядов для составной частицы произвольной структуры. Тогда локальная энергия частицы, находящейся на позиции $k \in K$ имеет вид: $E_k(\omega) = \sum_{i=1}^{|U|} \varepsilon_i(\omega_k) u_i(\omega, k)$.

Активность частиц задается с помощью конечного множества V преобразований вида $v : \Omega \times K \rightarrow \Omega$, которые осуществляет локальное изменение конфигурации и реализует активность частицы. С каждой частицей $\alpha \in A$ связан неотрицательный вектор преобразований $\varphi(\alpha) = (\varphi_1, \dots, \varphi_{|V|}) \in \mathbf{R}^{|V|}$, а с каждой n -местной связкой $\gamma \in \Gamma$ связана функция $\varphi_\gamma : \mathbf{R}^{n|V|} \mapsto \mathbf{R}$ агрегации n векторов активности частиц, находящихся в связке, что позволяет вычислять вектор активности для любой составной частицы. Тогда локальное преобразование $F_k(\omega)$, связанное с частицей на позиции $k \in K$ заключается в выполнении, одного из преобразований с вероятностью, пропорциональной величине соответствующего элемента вектора активности, т. е. $F_k(\omega) = v_i(\omega, k)$ с вероятностью $\varphi_i(\omega_k) / \sum_{j=1}^{|V|} \varphi_j(\omega_k)$, $i = \overline{1, |V|}$.

Моделирование универсальных внутриклеточных процессов. Рассмотрим один из возможных способов моделирования механизма экспрессии генов в рамках предложенной модели. Как отмечалось выше, механизм экспрессии генов является аналогом универсального вычислительного процесса в компьютере, поскольку позволяет запускать любые внутриклеточные процессы по соответствующим РНК-кодам путем синтеза функциональных белков. Рибосома – сложный молекулярный комплекс, лежащий в основе механизма экспрессии генов, также состоит из белков, РНК-коды которых наряду с другими белками хранятся в геноме.

Для моделирования механизма экспрессии генов необходимо предусмотреть способ кодирования функциональности частицы, а также смоделировать работу рибосомы в виде активности частицы.

Функциональность любой частицы определяется соответствующими векторами активности и зарядов частицы, которые в свою очередь вычисляются по структуре частицы. Структура любой частицы однозначно определяется символьной записью, например, частицы, структура которой показана на рис. 4, соответствует запись: " $\gamma_1[\beta_1, \gamma_1[\beta_2, \beta_3]]$ ". Такое представление частицы – это конечная последовательность символов из конечного алфавита, однозначно определяющая активность частицы.

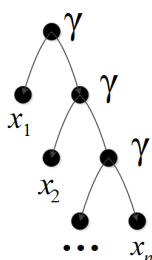


РИС. 4. Структурное представление любой частицы

Любую конечную последовательность x_1, \dots, x_n символов из конечного алфавита можно закодировать с помощью древовидной частицы специальной структуры, состоящей из специальных базовых частиц и связей. Аналогичным образом в живой природе аминокислотные последовательности белков, состоящие из 20 различных аминокислот, кодируются с помощью РНК-кодов, состоящих из нуклеотидов 4 типов (аденин, гуанин, цитозин, тимин).

Рибосома – это самый сложный молекулярный комплекс, строящий белки по РНК-кодам, молекулярный принцип работы которого остается до конца неясным. Тем не менее, с точки зрения функциональности рибосома ставит в соответствие триплету нуклеотидов одно аминокислотное основание. Рибосома работает настолько надежно, что можно утверждать о реализации ею детерминированной функции, которая одинаковым РНК-кодам всегда ставит в соответствие одинаковые аминокислотные последовательности. Эта функция универсальна для всех живых организмов и называется генетическим кодом. Универсальный генетический код имеет вид таблицы, с помощью которой, каждому из 64 возможных триплетов нуклеотидов ставится в соответствие один аминокислот. Такое преобразование несложно реализовать алгоритмически и смоделировать в виде активности некоторой частицы.

Заключение. В работе исследована роль понятие универсальности при управлении внутриклеточными процессами. Выяснилось, что аналогично цифровым вычислениям, понятие универсальности имеет глубокое значение для реализации внутриклеточных процессов в клетке. В работе предложен способ моделирования универсальных внутриклеточных процессов, таких как экспрессия генов с помощью ансамблей активных заряженных частиц.

Б.О. Білецький, А.М. Гупал

УНИВЕРСАЛЬНІСТЬ У МОДЕЛЮВАННІ ВНУТРІШНЬОКЛІТИННИХ ПРОЦЕСІВ

Обговорюється важливість поняття універсальності для моделювання внутрішньо-клітинних процесів. Наводяться паралелі між внутрішньоклітинними процесами та обчисленнями на ЕОМ, при цьому сама клітина порівнюється з універсальним обчислювальним пристроєм.

B.O. Biletskyy, A.M. Gupal

NOTION OF UNIVERSALITY IN THE INTRACELLULAR PROCESS MODELING

The paper deals with the importance of the notion of universality for intracellular process modeling. Intracellular processes coordinated by proteins are compared with digital computations and the cell itself is compared with a universal computer.

1. *Turing A.* On computable numbers, with an application to the Entscheidungsproblem // Proceedings of the London Mathematical Society. – 1937. – N 42. – P. 230 – 265.
2. *Schrödinger E.* What Is Life? // Cambridge University Press, 1944. – 194 p.
3. *Beletskiy B.A.* Collective Processes Modeling by Means of Particles with Complex Structure // Journal of Automation and Information Sciences. – 2011. – N 43. – P. 34 – 43.

Получено 15.01.2015

Об авторах:

Белецкий Борис Александрович,

кандидат физико-математических наук,
Института кибернетики имени В.М. Глушкова НАН Украины,
E-mail: borys.biletskyy@gmail.com

Гупал Анатолий Михайлович,

доктор физико-математических наук, член-корреспондент НАН Украины,
профессор, заведующий отделом
Института кибернетики имени В.М. Глушкова НАН Украины.