

УДК 004.021

## ИМИТАЦИЯ МОДЕЛИ ЗАРОЖДЕНИЯ ЖИЗНИ

*Протьюко М.А., студент гр.050502*

*Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь*

*Борисенко О.Ф. – канд. физ.-мат. наук*

**Аннотация.** Данная статья содержит описание «жизни» с точки зрения биоинформатики, а также возможный способ симуляции модели развития систем, способных этому определению соответствовать.

**Ключевые слова.** Континуальные модели, клеточный автомат, COMETS

### **Введение**

Среди огромного разнообразия гипотез о возникновении жизни, особо выделяется та, что утверждает, будто бы жизнь развилась постепенно, все больше и больше усложняя цепочки элементов, которые позже образовали РНК и белки, а еще позже, первую функциональную клетку (или же клеточную структуру) [1]. Доказательством теорий возникновения жизни занимается синтетическая биология, основная цель которой - в лабораторных условиях воспроизвести биологическое поведение неестественных биологических систем. Аналогично синтетической биологии, было создано множество программных симуляций, позволяющих подробно описать законы жизни (COMETS, VacArena, Matnet, 3DdFBA и др.) [2], настолько подробно, что возможно имитировать поведение различных микроорганизмов, процесс мутации и эволюции, а также многие химические реакции, описывающие метаболизм.

Свойство, которое позволяет живому организму изменять свое поведение и структуру называется адаптацией, а в долгосрочном плане – эволюцией. Если разложить понятие «жизнь» на элементарные составляющие, сможем ли мы создать некое подобие эволюционного процесса, только в своей «неестественной системе»?

Цель данной статьи – определить такие элементы и свойства живого, которые возможно реализовать вычислительными средствами, подобно [2].

### **Природа жизни**

Определим желаемый результат – выход данной симуляции, или же, «жизнь».

Согласно [3]:

«Биологический организм – это естественно возникающее, самовоспроизводящееся устройство, которое действует за счет управляемого превращения вещества, энергии и информации.» - далее, в данной работе будет использоваться именно это определение.

Живая клетка обладает следующими характеристиками:

- Метаболизм
- Энергия
- Клеточный рост, клеточное деление, синтез ДНК
- Транскрипция
- Синтез белков
- Пассивный транспорт
- Клеточный транспорт и транспортные механизмы
- Защита клеток, клеточная смерть и старение

Согласно [3], «можно предположить, что минимальный организм должен обладать способностью к воспроизведению, но не обязательно должен конкурировать в способности роста и размножения с другими организмами. Можно предположить, что минимальный организм должен усваивать питательную среду, обеспечивая организму биосинтез, а также обеспечивать ответ на стресс, в том числе и восстановление поврежденной ДНК.»

Т.е., необходимо задать такой набор правил, чтобы возможно было с их помощью описать метаболизм, а также набор катализаторов данного процесса (или же, аналогию с ДНК, что позволяет хранить информацию в системе).

### **Типы моделей**

Рассмотрим общие математические модели, описывающие жизнь на примере биопленки (biofilm dynamics [4], биопленка – сообщество микроорганизмов, адаптирующееся к внешним условиям, причем микроорганизмы по отдельности могут это свойство не иметь).

Существуют следующие модели имитации сообщества микроорганизмов:

- Континуальные модели (continuum models)
- Модели IbM (individual based - индивидуальные)
- Модели CA (клеточный автомат)
- Гибридные

Как показали источники [3-4] использование вышеописанных моделей обладает следующими свойствами:

Континуальная модель позволяет имитировать физические законы, но не организм, способный адаптироваться (рассматривают микроорганизм как непрерывный материал и граничные условия, по которым на основе формул вязкости и диффузии, уравнения массопереноса и т.п. рассчитывается конечное состояние, представляемое набором численных параметров).

Индивидуальный подход (IbM) крайне тяжело оптимизировать и посчитать. По сути, IbM представляют собой список из одного и того же уравнения с разными параметрами, такими как скорость роста (условия роста) и распространение биомассы. Для каждого конкретного вещества для описания его реакций (а следовательно, и жизненного цикла) необходима своя формула, и свой расчет. Таких сфер (микроорганизм в данной модели – твердая сфера) в одной симуляции минимум  $10^6$  на единицу поверхности (по аналогии с квадратным сантиметром). Причем данная модель не предусматривает эволюцию, а больше подобна на игру «Жизнь» Конвея, позволяя из заданных начальных значений составить топологию.

Клеточный автомат (по определению из [3]) подобен объединению идеи генетических алгоритмов, описывающих имитацию конкуренции (Роевой интеллект, симуляция «хищник – жертва»). Основная проблема данного подхода – необходимо выбрать следствия, относительно которых будет строиться модель. Следствия – это набор простых правил (питание, гибель, размножение), взятые на основе биологических принципов. CA способны моделировать многие простые взаимодействия, но для более комплексных (формирование биопленки, образование матрицы EPS) они, как показала практика, не пригодны.

Из вышеописанного сделаем вывод, что для имитации процесса эволюции жизни необходимы гибридные модели.

Далее, опишем свойства гибридной модели, разделив процесс на несколько уровней, подобно уровням организации живого.

### **«Физика жизни»**

Начнем с выбора такой системы простейших правил, которая позволяет рассчитать (1-4). В данном случае, определим, что будет в данной системе истинной, а что будет ложью (система аксиом [5])

Попробуем описать правила этого простейшего клеточного автомата, подобно правилам «Жизни» Конвея.

Элемент (на данном уровне – клетка, точка на пространстве, слот в «хромосоме» генетического алгоритма и т.д.) обладает двумя параметрами: скоростью передвижения (или же энергией) и положением в пространстве (описываемое неким радиусом действия элемента). Свойство энергии – стремление к энтропии, или же, равномерному ее распределению по всей симуляции. Т.е., элемент с большим Q стремится к элементу с малым Q, образуя тем самым связь. Данную связь можно разорвать другим элементом.

Т.е., поведение всей симуляции на данном уровне определяется двумя константами и двумя переменными: энергия связи, энергия разрыва – константы; энергия элемента, радиус действия элемента в некой системе координат - переменные. В системе задается время  $t$ , общее для всех уровней.

### **«Химия жизни»**

Большая часть уравнений, описанных в данном разделе, рассчитывается с помощью матрицы стехиометрических коэффициентов. Разумеется, данная матрица должна присутствовать для каждого метаболита.

В качестве метаболита в имитационной модели может использоваться любой набор элементов «физического» уровня, стехиометрические коэффициенты которого будут посчитанные на основе формул, позволяющих определить характеристики живой клетки [3] (к примеру, матрица стехиометрических коэффициентов, описывающая элемент, подобный клеточной стенке бактерий и т.д.).

Уравнения, позволяющие описать усвоение питательной среды и биосинтез:

$$\frac{\partial B^\alpha}{\partial t} = \vartheta^\alpha B^\alpha \quad (1)$$

Где  $B^\alpha$  – биомасса вида  $\alpha$ ,  $\vartheta^\alpha$  – скорость роста, посчитанная через FBA (Flux Balance Analysis),  $\alpha$  – количество элементов некоего вида.

$$\frac{\partial Q^i}{\partial t} = \sum_{\alpha} \vartheta_i^\alpha B^\alpha \quad (2)$$

Где  $Q^i$  – объем внешнего метаболита (abundance of external metabolite),  $\vartheta_i^\alpha$  – скорость обмена метаболитом  $i$  в виде  $\alpha$ .

Уравнения, позволяющие определить самовоспроизведение [6], это возможные зависимости между внешним и внутренним метаболитом: линейная функция концентрации, монотонная (Михаэль-Ментен) функция, или псевдо-монотонная функция.

Рассмотрим монотонную модель, описывающую отношения между ростом и концентрацией питательного субстрата (в случае данной модели – взаимоотношение между связанными элементами (уже на «химическом» уровне) и свободными).

$$\mu = \mu_{max} \left[ \frac{S_i}{S_i + K_i} \right] \quad (3)$$

Где  $\mu$  – скорость роста элементов  $\alpha$ ,  $S_i$  – концентрация метаболита  $i$ .  $\mu_{max}$  – максимально допустимая скорость роста для данного  $\alpha$  (выбираемая константа),  $K_i$  – насыщение  $i$  (в данной системе насыщением можно назвать доступную энергию метаболита. Расчет данной энергии  $Q$  (из уровня «физиологии» жизни производится расчет  $K_i$  по формуле, аналогичной формуле расчета сатурации кислорода. Данный расчет также зависит от выбора констант (уже для «химического» уровня) – для разделения метаболитов на активные, катализаторы и т.д.)

На основе модели Верхулста, или же метода идентификации логической динамики и жизненного цикла метаболита:

$$X(t) = \frac{X_i X_0 e^{\mu t}}{X_i - X_0 + X_0 e^{\mu t}} \quad (4)$$

Где  $X$  – концентрация выбранного метаболита  $i$ ,  $X_i$  – его максимальная концентрация,  $X_0$  – начальная концентрация метаболита  $i$ ,  $t$  – время,  $\mu$  – скорость роста элемента из (3).

Формулы (3-4) были выбраны поскольку для них существует статистика из исследований роста и эволюции микроорганизмов ([6] и [4])

Формулы (1-2) были выбраны, поскольку они хорошо показывают себя в симуляции COMETS.

Матрицу стехиометрических коэффициентов можно рассчитать на основе FBA (Flux Balance Analysis), которые позволяют, на основе некоторых ограничений предсказывать функцию или фенотип организма на основе его метаболизма [2]. Расчет данной функции также требует задания специфических формул и констант.

### Информация

Согласно [7], в качестве параметра передачи информации выступает соответствие между одной группой элементов другой. Т.е., в данной модели должен присутствовать код, где изначальным алфавитом будет выступать набор неких «органических» метаболитов, а кодовым алфавитом – набор метаболитов-катализаторов. Данный процесс позволяет переходить из одной группы элементов к другой. Причем, данный код обладает избыточностью.

Избыточность кода можно получить путем сохранения соотношения между емкостью кода и количеством символов источника, подобно соотношению ДНК – аминокислоты.

К примеру, для длины кода 2 с основанием 4 получаем емкость 16, причем не кодирующих элементов может быть от 1 до 6. Оставшиеся последовательности кодируют 5 изначальных элементов.

Причем не столь важно, чтобы каждый метаболит из множества «органических» имел свою кодовую последовательность. Достаточно описания минимальных метаболитов, способных посредством дальнейших взаимодействий друг с другом создать необходимый набор (данный набор описывается FBA на «химическом» уровне).

### Заключение

Данная модель требует большого количества параметров, требующих выбора (задание начальных параметров). Выбор параметров возможно оптимизировать. Поскольку имеется огромное количество данных, описывающих взаимодействия (1-4), можно использовать аналогию с

помощью нейронной сети, которая, обучаясь на выборке из работ, подобных [4] будет способна распознавать потенциально выигрышнее последовательности, задавая нужные входные параметры «физического» уровня и константы «химического». Т.е., цель данной нейронной сети - найти такой набор констант, чтобы принципы и формулы, описанные на каждом уровне, были верными.

Полученные решения теоретически должны соответствовать определению жизни.

Но стоит учитывать, что формулы (1-4) – лишь немногие из возможных разновидностей. Т.е., решения может и не быть.

Основные отличия данной модели от генетических алгоритмов и прочих симуляций органических процессов:

Элементы конечны – их количество остается неизменным с начала симуляции (т.е., энергия не возникает из ниоткуда и не исчезает в никуда).

Законы более высокого уровня организации выводимы из законов низкого уровня (т.е., данная модель – не набор следствий, как в классическом клеточном автомате биопленок, а набор причин)

**Список использованных источников:**

1. Беннер С.А., Ким Х.Дж., Ян З.: Подготовка сцены: история, химия и геобиология, стоящие за РНК. Колд Спринг Харб Перспектива Биол. 2012, 4: a003541-10.1101
2. Ilija Dukovski, Djordje Bajić, Jeremy M Chacón u др., *Computation Of Microbial Ecosystems in Time and Space (COMETS): An open source collaborative platform for modeling ecosystems metabolism* – [эл.источник] – URL - <https://arxiv.org/ftp/arxiv/papers/2009/2009.01734.pdf>
3. Леск А. Введение в биоинформатику / под ред. Миронова А.А, Шведаса В.К. // Москва, БИНОМ.Лаборатория знаний 2009 – 326 с.
4. David Rodriguez Espeso. *MODELING AND SIMULATION OF BACTERIAL BIOFILMS / Thesis Doctoral, DEPARTAMENTO DE CIENCIA E INGENIERÍA DE MATERIALES E INGENIERÍA QUÍMICA, Leganés, Junio de 2013 – 245.*
5. Протько М.А. Формализация и исследование замкнутых систем. 59 конференция студентов, магистрантов и аспирантов БГУИР, 2023, 3 с.
6. Siti suhana Jamaian *The Extended Monod Model for Microalgae Growth and Nutrient Uptake in Different Wastewaters // Universiti Tun Hussein Onn Malaysia International Journal of Engineering & Technology– 2018. 5 p.*
7. Протько М.А. Алгоритм кодирования процесса трансляции белков в клетке. Международный научно-технический семинар «Технологии передачи и обработки информации». БГУИР. 2023,6 с.

UDC 004.021

## DESCRIPTION OF THE MODEL OF THE ORIGIN OF LIFE

*Protsko M.A.*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*Borisenko O.F. – PhD in Physics and Mathematics*

**Annotation.** This article contains a description of "life" from the point of view of bioinformatics, as well as a possible way to simulate a model of the development of systems capable of meeting this definition.

**Keywords.** Continuum models, cellular automata, COMETS