

МОДЕЛИРОВАНИЕ РАБОТЫ ДНК

Р.Г. Мустафин

АННОТАЦИЯ

В работе рассматриваются возможные алгоритмы формирования структуры биологических объектов на основе генетической информации, записанной в ДНК каждой клетки. Рассмотрены алгоритмы клеточных автоматов, волновые алгоритмы, позволяющие на основе малого объема информации построить неограниченно сложные структуры биологических объектов.

Ключевые слова: ДНК, клеточный автомат, волновой алгоритм, компьютерное моделирование.

ABSTRACT

The paper considers possible algorithms for the formation of the structure of biological objects based on the genetic information recorded in the DNA of each cell. Algorithms of cellular automata and wave algorithms are considered, that make it possible to build unlimited complex structures based on a small amount of information.

Keywords:

DNA, cellular automaton, wave algorithm, computer simulation.

Введение

Среди методов моделирования работы биологических объектов существует модель «клеточного автомата», которая разрабатывалась начиная с середины 20 века (Станислав Улам, Джон фон Нейман, Норберт Винер, Артуро Розенблют [1-3]. «The cellular automaton consists of a line of cells, each colored either black or white. At every step there is then a definite rule that determines the color of a given cell from the color of that cell and its immediate left and right neighbors on the step before» [4].

Подобные модели клеточных автоматов можно применить для моделирования работы генетического механизма, алгоритм которого записан в ДНК, алгоритм формирования формы тела: от единственной клетки, путем деления и специализации клеток создание сложной структуры тела.

Рассмотрим простейшую модель: есть некие структуры (клетки), каждая из которых имеет полный набор программы (ДНК) действий

(специализации). Задача – создать такой набор программ действий, в результате работы которой получим из неспециализированных клеток – некую структуру (организм). Таким образом цель работы – путем моделирования продемонстрировать возможность создания сложной структуры используя достаточно малый объем программы. То есть разрешить противоречие: объема программы ДНК [5] недостаточно для описания работы, специализации каждой клетки организма (число клеток организма [6] гораздо больше информации, записанной в ДНК). Кроме того, моделирование может прояснить принципы специализации клеток в процессе роста организма.

Методика эксперимента

Моделируем специализацию клеток, в зависимости от записей в ДНК: каждая клетка имеет одинаковую программу, записанную в ДНК. Для простоты рассмотрим специализацию клеток из заданного массива неспециализированных, универсальных клеток. Специализация клетки зависит от программы, которая для выбора дальнейшей специализации данной клетки рассматривает её ближайшее окружение: для одномерного случая – два соседа, двумерного (квадратная решетка на плоскости) – четыре, трехмерного – 6 (для прямоугольной решетки и для ближайших соседей), возможен учёт более дальних соседей.

Что хотим получить?

- Скин, поверхность, на границе между окружающей средой и внутренними областями, толщиной в один слой клеток. Клетка на границе видит внешнюю среду (соседние клетки представляют собой внешнюю среду) и специализируется в поверхностные клетки, создавая поверхность (Скин) толщиной в одну клетку.
- Сетевую конструкцию (модель нервной системы, кровеносной системы): от единого центра – равномерно заполняет древовидной структурой. Нерв выделяет вокруг себя некие маркеры, где их нет – туда разрешено расти нервной системе. Таким образом избегаются пересечения с соседними нервными волокнами и создается равномерное заполнение свободного пространства. Алгоритм аналогичен волновым методам поиска кратчайшего пути на планарном графе от стартовой ячейки к конечной ячейке [7], от стартовой ячейки порождается шаг в соседнюю ячейку, при этом проверяется, проходима ли она, и не принадлежит ли ранее меченной в пути ячейке.

Обсуждение результатов

Созданная компьютерная программа моделирования работает следующим образом: в начале задан массив не специализированных клеток, окружающую среду и начальный массив нейронов (Рис.1).



Рис.1. Начальное состояние. Желтые пиксели – не специализированные клетки, черные пиксели – поверхность, на которой закреплены клетки (и откуда идет питание клеток), белые пиксели – окружающая среда, красные пиксели – массив нейронов.

Программа последовательно перебирает клетки (пиксели), и обрабатывает их (создает специализацию, если есть необходимость).

При такой последовательной обработке клеток (пикселей) – результат зависит от последовательности обхода клеток (особенно для нейронов, которые стремятся заполнить равномерно всю предоставленную площадь) (Рис.2).

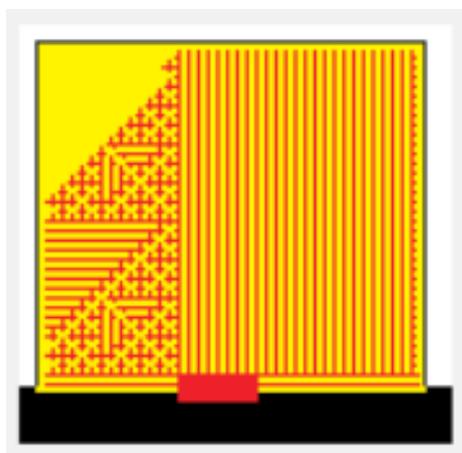


Рис.2. Состояние после 100 итераций (обходов всех пикселей). Серые пиксели по краям – созданный Скин (поверхность толщиной один пиксел). Почти вся свободная поверхность занята нейронами, которые не пересекаются друг с другом, и между которыми зазор в один пиксел. Левый верхний угол не заполнен – не хватило числа итераций.

Дальнейшее развитие программы – специализация клеток со случайной последовательностью (каждая клетка специализируется в случайный момент времени). При добавлении некоторой случайности картина становится более однородной (Рис.3).

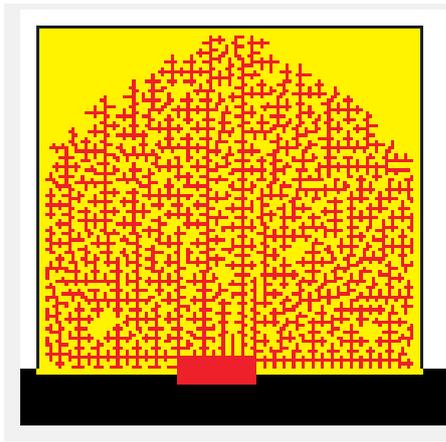


Рис.3. Состояние после 1000 итераций (обходов всех пикселей). Серые пиксели по краям – созданный Скин (поверхность толщиной один пиксел). Почти вся свободная поверхность занята нейронами, которые не пересекаются друг с другом, и между которыми зазор в один пиксел. Левый верхний и правый угол не заполнены – не хватило числа итераций.

Двумерная модель специализации клеток накладывает некоторые ограничения (например, для развития нейронной системы на поверхности одна нить нейронов перекрывает путь всем остальным), однако проста для реализации. Кроме того, клетки представлены квадратами, в реальности клетки имеют произвольную форму. Однако все эти ограничения все равно позволяют продемонстрировать возможность малым объемом программы (малым участком ДНК) занимать неограниченный участок (край участка – для Скина, весь участок – для нейронов).

Дальнейшее возможное развитие программы моделирования: а) включение деления клеток (что в данной статье не рассматривается) плюс специализация клеток, б) попытка смоделировать создание формы: когда из первоначальной клетки, в процессе деления и специализации, возникает форма тела.

Гораздо большую сложность представляет моделирование замкнутых сетевых структур, типа кровеносной системы. В отличие от нейронной сети (структура которой представляет собой дерево, однонаправленное, растущее от основания и старающееся занять весь доступный участок) кровеносная система имеет две транспортные ветки: поставляющую кровь (артериальная система) и возвращающая кровь обратно (венозную систему). При этом кровеносная система также стремится занять весь доступный участок тела. Трудность представляет то, что кровеносная система замыкается (артериальная система на венозную) на самых концевых участках (в системах капилляров).

Возможный алгоритм роста кровеносной системы может быть такой: подобно нервной системе, артериальная система растет как расходящееся дерево, занимая весь доступный участок тела, и в тех участках, где

содержание кислорода в крови становится малым, артериальная система замыкается на венозную.

Заключение

Таким образом, на простых моделях (Скина, поверхности толщиной в одну клетку, нейронной сети, линейные нейроны, выходящие из одной области и занимающие, охватывающие весь доступный участок) показана возможность создания сложной структуры используя достаточно малый объем программы. Такое моделирование работы генетической системы ДНК может помочь в понимании алгоритмов работы ДНК в живых системах, может показать путь, алгоритмы создания сложных, самоорганизующихся объектов, состоящих из множества одинаковых ячеек, взаимодействующих друг с другом.

Литература

1. Pickover, Clifford A., Pickover, Clifford A. The Math Book: From Pythagoras to the 57th Dimension, 250 Milestones in the History of Mathematics. — Sterling Publishing Company, Inc, 2009. — ISBN 978-1402757969.
2. A.G. Hoekstra, J. Kroc, P. Sloot. Simulating complex systems by cellular automata. Springer, 2010. ISBN 978-3-642-12202-6.
3. Liana Manukyan, Sophie A. Montandon, Anamarija Fofonjka, Stanislav Smirnov & Michel C. Milinkovitch. A living mesoscopic cellular automaton made of skin scales // Nature. — 2017. — Vol. 544. — P. 173–179. — DOI:10.1038/nature22031.
4. Wolfram S. A new kind of science. Wolfram media, 2002] (<https://www.wolframscience.com/nks/p23--how-do-simple-programs-behave/>).
5. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. (англ.) // Nature. — 2004. — Vol. 431, no. 7011. — P. 931—945. — DOI:10.1038/nature03001 (The current genome sequence (Build 35) contains 2.85 billion nucleotides interrupted by only 341 gaps).
6. Sender R, Fuchs S, Milo R (2016) Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLOS Biology 14(8): e1002533. <https://doi.org/10.1371/journal.pbio.1002533>.
<https://www.biorxiv.org/content/10.1101/036103v1> (estimates to reach a total of $3.0 \cdot 10^{13}$ human cells in the 70 kg “reference man”).
7. Moore E. F. The shortest path through a maze // Proceedings of an International Symposium on the Theory of Switching (Cambridge, Massachusetts, 2–5 April

1957) — Harvard University Press, 1959. — Vol. 2. — P. 285–292. — 345 p. —
(Annals of the Computation Laboratory of Harvard University; Vol. 30).