

УДК 548.31

АЛГОРИТМИЧЕСКАЯ КРИСТАЛЛОХИМИЯ: МЕТОД КЛЕТОЧНЫХ АВТОМАТОВ

© 2012 г. С. В. Кривовичев

Санкт-Петербургский государственный университет

Отдел исследования микро- и нанопористых структур, КНЦ РАН, Апатиты

E-mail: skrivovi@mail.ru

Поступила в редакцию 16.02.2011 г.

Атомно-молекулярные механизмы роста кристаллов могут быть смоделированы на основе кристаллохимической информации с использованием клеточных автоматов. В частности, формирование гетерополиэдрических слоистых комплексов в селенатах уранила может быть смоделировано с использованием одномерного трехцветного клеточного автомата. Использование теории вычислений, в частности теории автоматов, в кристаллографии позволяет интерпретировать рост кристалла как вычислительный процесс (выполнение алгоритма или программы с конечным числом шагов).

ВВЕДЕНИЕ

Одной из фундаментальных проблем кристаллографии является проблема образования трехмерной упорядоченной периодической структуры, которую Н.В. Белов называл проблемой “501-го атома” [1]. В 1970–1980-е гг. Н.В. Белов поставил эту проблему перед сотрудниками математической школы Б.Н. Делоне, что привело к созданию оригинальной теории локального критерия правильности точечных систем [2–6], согласно которой правильность системы точек задается равенством локального окружения точки остальными точками системы внутри некоторой области. При всей удивительной красоте и математической строгости эта теория не нашла должного признания среди “практических” кристаллохимиков, привыкших к реальным структурам, состоящим не из математических точек, а из физически реальных атомов и молекул. Можно сказать, что между математической теорией формирования правильных систем и эмпирическими данными о многообразии известных кристаллических структур возник некоторый разрыв. С другой стороны, налицо отсутствие какой-либо теории, связывающей кристаллохимию с теорией роста кристаллов. В общепринятых моделях роста кристаллов (будь то теория Косселя в традиционном [7] и современном [8] вариантах или модель роста на дислокациях [9]) кристалл рассматривается как построенный из абстрактных частиц кубической формы, независимо от того, является ли он кристаллом поваренной соли или сложного цеолита, содержащего внутри элементарной ячейки сотни симметрично неэквивалентных

атомов. Вместе с тем для некоторых групп соединений (в частности, для промышленно важных соединений с микропористыми структурами) в последнее время достигнут существенный прогресс в понимании атомно-молекулярных механизмов роста их кристаллических образований [10, 11]. При этом первостепенную роль играют данные о кристаллохимических особенностях соединения, присутствия в его структуре тех или иных фундаментальных строительных блоков и т.д.

Настоящая работа посвящена развитию алгоритмической кристаллохимии, в которой построение той или иной кристаллической структуры рассматривается как следствие выполнения того или иного алгоритма или программы. Вследствие периодичности кристаллического вещества алгоритм его построения должен быть по существу циклическим, т.е. включать в себя выполнение практически бесконечного числа одной и той же программы, состоящей из конечного числа операций присоединения строительных частиц — атомов, молекул, кластеров, наноблоков (в случае неклассической кристаллизации [12] и т.п.). Каждая из этих операций может быть описана на языке кристаллохимии. Таким образом, расшифровка алгоритма роста структуры заключается в выделении конструктивных этапов присоединения тех или иных строительных блоков. В настоящей работе развивается метод моделирования процессов роста кристаллических структур при помощи клеточных автоматов как частного случая конечных детерминированных автоматов.

ТЕОРИЯ КОНЕЧНЫХ АВТОМАТОВ: ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

В классической теории вычислений [13] конечный автомат определяется как набор $\{Q, \Sigma, \delta, q_0\}$ из следующих основных элементов:

- множество $Q = \{q_0, q_1, q_2, \dots, q_n\}$ из конечного числа n состояний автомата;
- конечного алфавита $\Sigma = \{a, b, c, \dots\}$ входных символов;
- отображения множества Q на себя через алфавит $\Sigma: Q \times \Sigma \rightarrow Q$ (правила перехода между состояниями автомата при вводе символов из алфавита Σ);
- начального состояния автомата q_0 .

Иногда к этому набору элементов добавляется подмножество состояний из Q , называемых принимающими состояниями (*accepting states*).

Простейшим примером конечного автомата является лифт. Множество Q его стационарных состояний соответствует множеству этажей в доме, набор входных символов – множеству кнопок на пульте внутри лифта. Правила перехода в этом случае просты: при нажатии кнопки с соответствующим номером лифт переходит в новое состояние (этаж), соответствующее номеру кнопки. Если в доме три этажа, то автомат, моделирующий лифт, состоит из трех состояний, трех входных символов и набором правил перехода между этажами. Такой автомат можно охарактеризовать диаграммой состояний (*state diagram*), представленной на рис. 1а. На диаграмме состояния изображаются в виде узлов с соответствующим номером, а переходы между ними изображаются стрелками, над которыми указаны символы переходов между этажами. Если лифт исправлен, то любая последовательность (строка) символов **1**, **2** и **3** будет соответствовать последовательному перемещению лифта между этажами. Иными словами автомат будет *принимать* такую последовательность. Допустим, лифт неисправен и не может переезжать с третьего этажа на первый и обратно без остановки на втором этаже (в этом случае, чтобы попасть с первого этажа на третий, надо сначала приехать на второй этаж). Диаграмма состояния автомата, соответствующего такой системе, изображена на рис. 1б. Очевидно, такой автомат будет принимать последовательность **1211232321**, но не примет последовательность **1232321321** ввиду наличия в ней элемента **13**, соответствующего прямому переезду с первого на третий этаж. Множество всех строк (слов), принимаемых некоторым конечным автоматом, образует *формальный язык* данного автомата, а правила перехода между состояниями соответствуют словообразующей *формальной грамматике* этого языка.

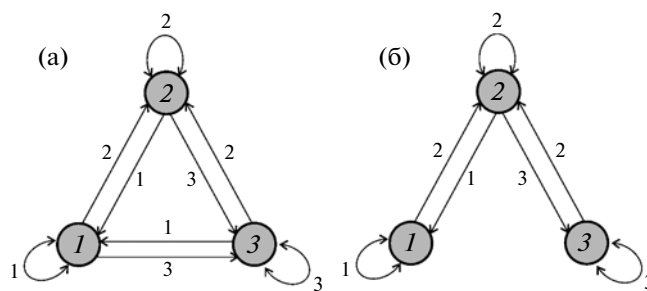


Рис. 1. Диаграмма состояния исправного (а) и неисправного (б) лифта в трехэтажном доме.

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ КЛЕТОЧНЫХ АВТОМАТОВ

Клеточные автоматы (КА) представляют собой детерминированные дискретные динамические системы, развивающиеся на основе локальных правил взаимодействия элементов системы [14–16]. С формальной точки зрения КА состоит из пяти компонентов

$$\text{КА} := \langle Z, S, N, f, B \rangle, \text{ где}$$

Z – дискретное пространство действия автомата, иногда называемое решеткой (*lattice*); Z может быть конечным или бесконечным; единицей Z является ячейка (*cell*); наиболее обычная решетка КА – плоская решетка с квадратными ячейками;

$S = \{0, 1, 2, \dots\}$ – конечный набор значений, которые могут принимать ячейки в процессе развития системы; в большинстве случаев, значения соответствуют определенным цветам: при $S = \{0, 1\}$ (т.е. когда мощность множества $|S| = 2$) имеет место двухцветный (бинарный) КА, ячейки которого могут принимать либо белый (0), либо черный (1) цвета;

$N = \{-k_1, -k_1 + 1, \dots, -1, 0, 1, \dots, k_2 - 1, k_2\}$ – окрестность (*neighborhood*) действия автомата – для одномерного автомата значение новой ячейки x_0 для момента времени $t = 1$ (новая строка) определяется значениями в момент времени $t = 0$ (текущая строка) самой ячейки x_0 , а также k_1 и k_2 ячеек с левой и правой стороны соответственно; для простейшего случая $k_1 = k_2 = 1$: окрестность симметрична, имеет радиус $k_0 = 1$ и состоит из трех ячеек ($k = 3$): x_{-1}, x_0, x_1 ;

f – локальная функция перехода (для одномерного автомата – правило построения новой строки, исходя из состояния текущей строки), определяемая окрестностью действия автомата; обычно записывается как набор правил вида $010 \rightarrow 1$ (для бинарного автомата с симметричной окрестностью радиуса $k_0 = 1$);

B – граничные условия (например, КА может быть определен на плоскости, на цилиндре, на поверхности Мебиуса и т.п.).

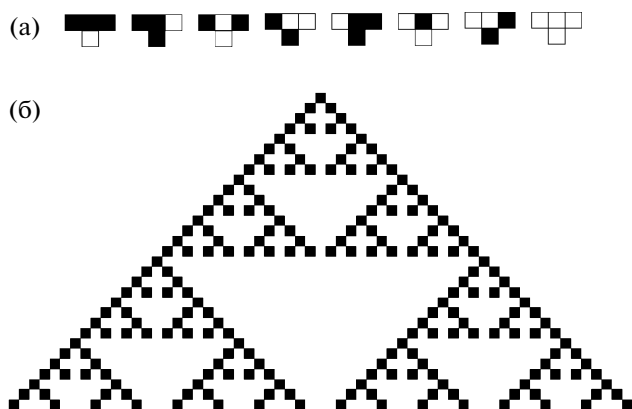


Рис. 2. Правила перехода автомата КА90 (а) и диаграмма его развития после 30 шагов (временных интервалов).

В качестве примера рассмотрим одномерный бинарный автомат, изображенный на рис. 2. В классификации С. Вольфрама этот автомат имеет

номер 90. Принимая значения 0 и 1, соответствующие белому и черному цвету ячейки, локальная функция перехода для этого автомата записывается как

$$f = \{111 \rightarrow 0, 110 \rightarrow 1, 101 \rightarrow 0, 100 \rightarrow 1, 011 \rightarrow 1, 010 \rightarrow 0, 001 \rightarrow 1, 000 \rightarrow 0\}.$$

При входной строке, содержащей одну черную ячейку, работа КА90 приводит к образованию разветвляющегося узора, подобного треугольнику Серпинского (двумерному аналогу множества Кантора).

ИСПОЛЬЗОВАНИЕ КЛЕТОЧНЫХ АВТОМАТОВ ДЛЯ ОПИСАНИЯ РОСТА КРИСТАЛЛИЧЕСКИХ СТРУКТУР

Клеточные автоматы представляют собой удобный абстрактный аппарат для моделирования роста кристаллических структур [17–20], в том числе для предсказания новых, ранее неизвестных структурных топологий [21]. На рис. 3 изображены три двумерные гетерополиэдриче-

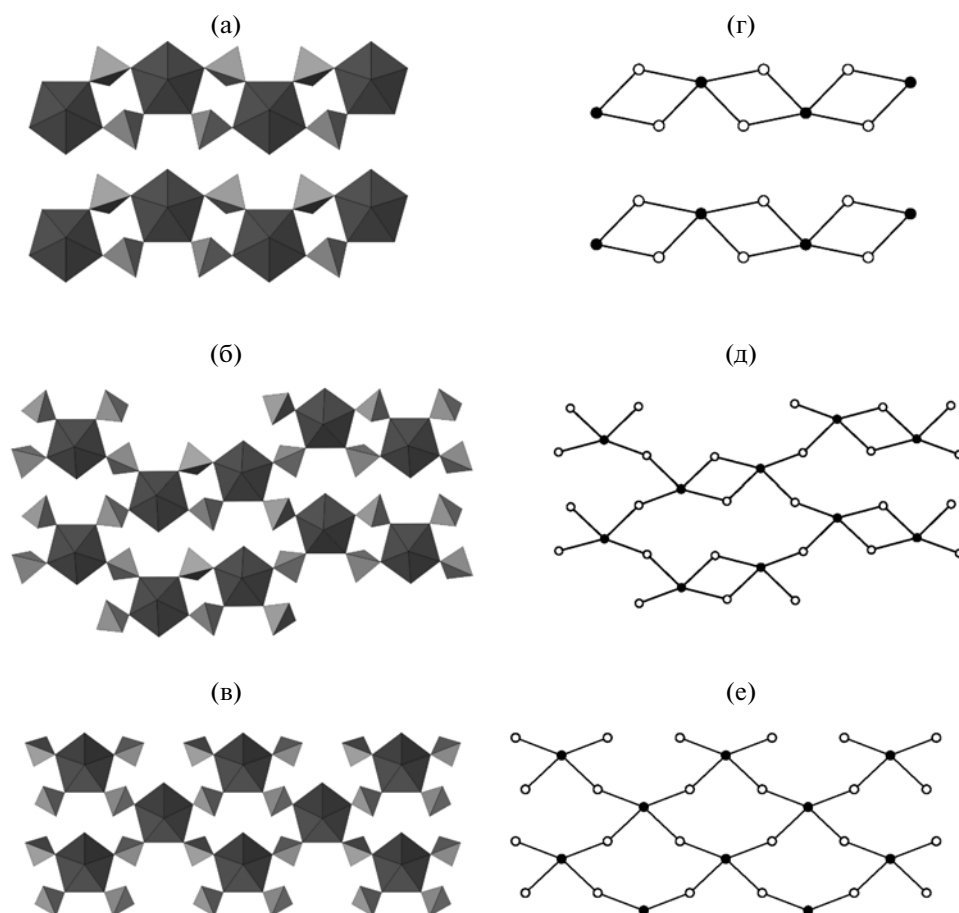


Рис. 3. Гетерополиэдрические слоистые структуры в селенатах, сульфатах и хроматах уранила (а, б, в) и их графы (г, д, е соответственно); темно-серый цвет – пентагональные бипирамиды уранила, светло-серый – тетраэдрические оксоанионы (а–в); черные и белые вершины символизируют уранильные и тетраэдрические комплексы (г–е); наличие ребра между вершинами в графах соответствует наличию общего мостикового атома кислорода между соответствующими координационными полиэдрами.

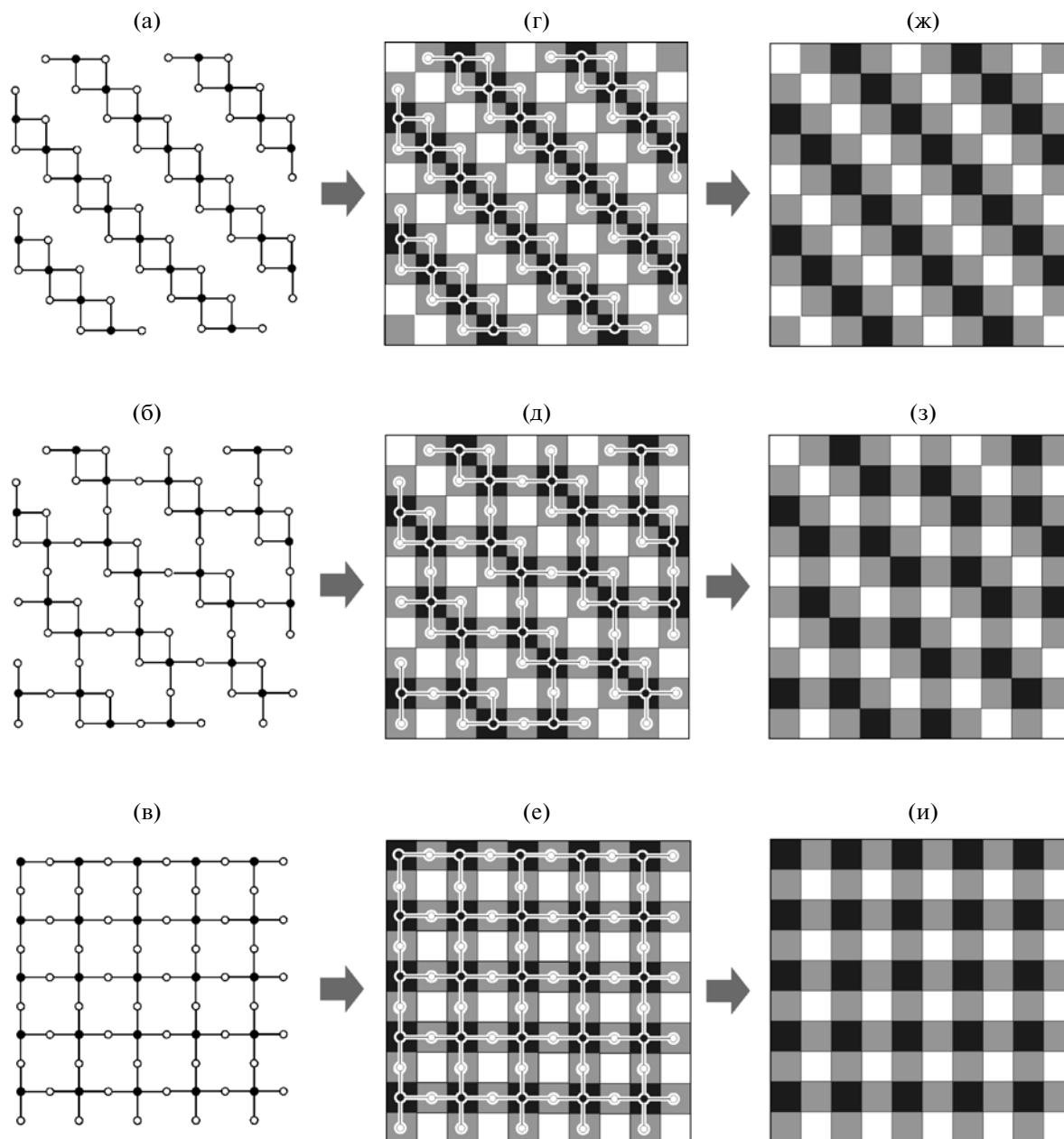


Рис. 4. Идеализированные варианты графов, показанных на рис. 3а–3в (а–в соответственно), схемы перехода от графов к ячейкам (г–е) и ячеистые структуры (ж–и).

ские структуры, обнаруженные в структурах солей шестивалентного урана (уранила) [22]. Пентагональные бипирамиды соответствуют координационным полиэдрам урана, тетраэдры – сульфатным, селенатным или хроматным тетраэдрическим оксоанионам TO_4 ($T = S, Se, Cr$). Топология объединения координационных полиэдров в гетерополиэдрическом комплексе может быть изображена при помощи его представления в виде графа [23, 24]. Графы комплексов, изображенных на рис. 4а–4в, представлены на рис. 4г–4е соответственно. При этом координационные

полиэдры урана и тетраэдрические оксоанионы символизируются черными и белыми узлами графа, а ребро между вершинами говорит о наличии между соответствующими координационными полиэдрами мостикового атома кислорода.

Для построения КА, моделирующего рост структур в солях уранила, преобразуем графы, изображенные на рис. 3г–3е, в графы, изображенные на рис. 4а–4в. Топология соответствующих графов остается инвариантной, иными словами эти графы изоморфны с комбинаторно-топологической точки зрения. При замене белых

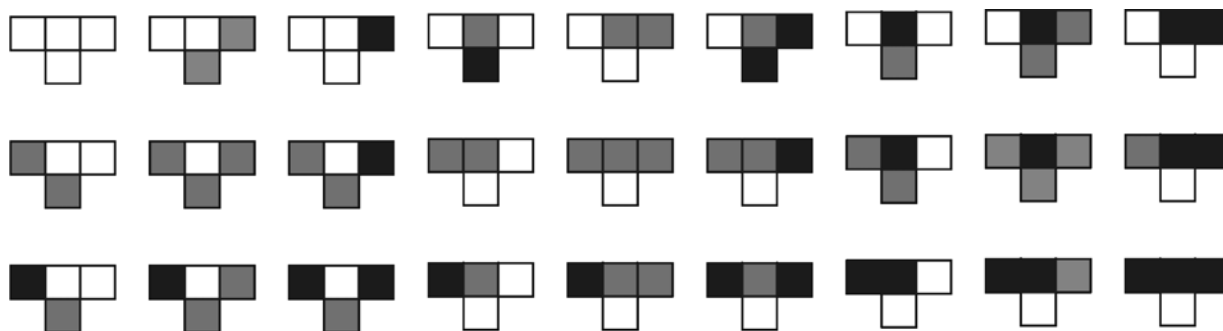


Рис. 5. Правила перехода для автомата, генерирующего ячеистые структуры, показанные на рис. 4ж–4и.

узлов серыми, а черных — черными ячейками, можно получить ячеистые структуры, изображенные на рис. 4ж–4и. Полученные структуры состоят из черных, серых и белых квадратных ячеек. Наличие между ячейками общей стороны соответствует наличию общего ребра между соответствующими черными и белыми вершинами графа. Анализ полученных структур показывает, что все три могут быть получены при работе тринарного (трехцветного) клеточного автомата, правила которого схематично изображены на рис. 5. Присвоив значения 0, 1 и 2 белым, серым и черным ячейкам, можно составить локальную функцию перехода в аналитическом виде.

Заметим, что для данного КА $|S| = 3$, а окрестность действия автомата симметрична и имеет радиус $k_0 = 1$ ($k = 3$). Следовательно, число правил перехода должно составлять $3^3 = 27$ (рис. 5). В аналитическом виде они записываются как $f = \{000 \rightarrow 0, 001 \rightarrow 1, 002 \rightarrow 0, 010 \rightarrow 2, 011 \rightarrow 0, 012 \rightarrow 2, 020 \rightarrow 1, 021 \rightarrow 1, 022 \rightarrow 0, 100 \rightarrow 1, 101 \rightarrow 1, 102 \rightarrow 1, 110 \rightarrow 0, 111 \rightarrow 0, 112 \rightarrow 0, 120 \rightarrow 1, 121 \rightarrow 1, 122 \rightarrow 0, 200 \rightarrow 1, 201 \rightarrow 1, 202 \rightarrow 1, 210 \rightarrow 0, 211 \rightarrow 1, 212 \rightarrow 0, 220 \rightarrow 0, 221 \rightarrow 0, 222 \rightarrow 0\}$.

Полученный трехцветный автомат с такими правилами перехода генерирует все ячеистые структуры, изображенные на рис. 4ж–4и, а следовательно, может быть использован для генерации соответствующих им графов и конкретных химических структур, моделями которых они являются. При этом образование конкретного комплекса будет определяться структурой начальной строки или, переводя на язык структурной химии, структурой зародыша.

Компьютерный эксперимент по изучению зависимости топологии комплекса от структуры начальной строки был проведен при помощи программы Mathematica 6.0 (встроенный модуль CellularAutomaton). В качестве начальной строки выбиралась бесконечная строка с периодически повторяющимся фрагментом. Некоторые результаты моделирования показаны на рис. 6. Как вид-

но из рисунка, цепочечная топология (рис. 3б, 3г, 3е) образуется при вводной строке с периодом **0121** (рис. 6а), слоистая на рис. 3б — при вводной строке с периодом **01012121** (рис. 6б), слоистая на рис. 3в — при вводной строке с периодом **12** (рис. 6в). При других вариантах вводных строк (рис. 6г–6е) образуются другие топологии, пока не встречающиеся в структурах солей уранила.

Таким образом, полученный КА моделирует процесс роста того или иного гетерополиэдрического комплекса в структурах сульфатов, селенатов и хроматов уранила.

Теория КА подразумевает формулировку некоторых понятий, имеющих прямое отношение к понятиям структурной химии, что делает КА еще более привлекательными с точки зрения моделирования химических структур. Одним из таких понятий является концепция “Эдемского сада” (Garden-of-Eden) как такой конфигурации, которая никогда не может возникнуть в результате работы КА (т.е. у этой конфигурации нет предшественников) [25]. С точки зрения химии это означает конфигурацию атомов и молекул, абсолютно невыгодную с энергетической точки зрения и, как следствие, абсолютно невозможную. Структура “Эдемского сада” подразумевает наличие некоторой минимальной конфигурации, появление которой невозможно. Иными словами каждый “Эдемский сад” должен содержать такую минимальную конфигурацию. В теории КА такая конфигурация называется “орфан” (от англ. *orphan* — “сирота”). В переводе на язык структурной химии орфан-структура — это тот элемент атомного кластера или группировки, который ее дестабилизирует и делает ее существование невозможной. Для рассматриваемого КА, симулирующего структуры селенатов уранила, орфан-конфигурациями будут последовательности **11** и **22**, а “Эдемским садом” — любая строка, их содержащая. При переходе к реальным химическим структурам эти последовательности будут соответствовать образованию связи между двумя тетраэдрами TO_4^{2-} (крайне невыгодной с энергетической точки зрения) или двумя пентагональными

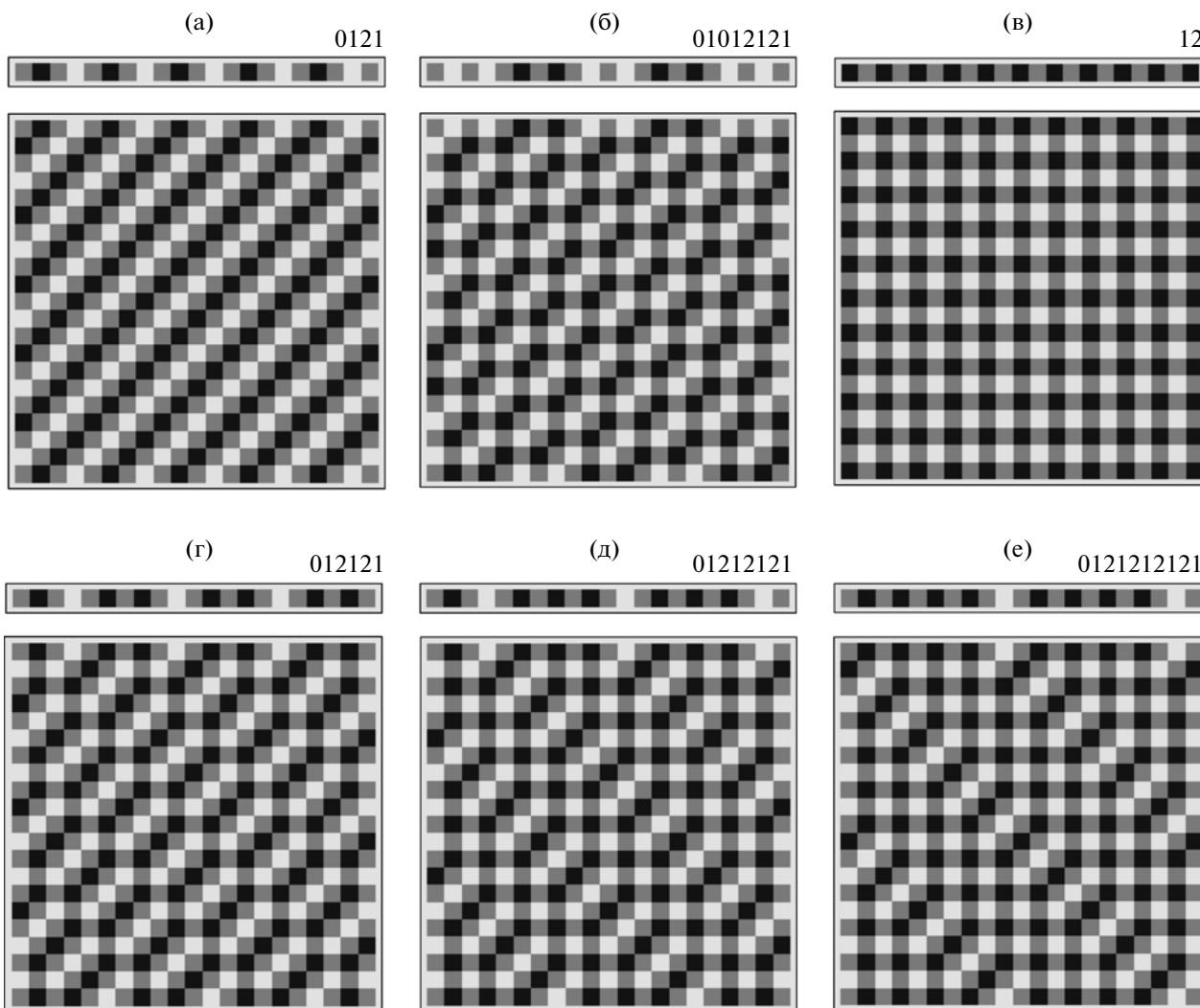


Рис. 6. Результаты компьютерного моделирования ячеистых структур на основе трехцветного клеточного автомата с правилами перехода, показанными на рис. 5, и бесконечной строкой с периодом (указан над каждой структурой в аналитическом и графическом видах).

бипирамидами уранила (что возможно, но не наблюдается в кислых средах, в которых растут кристаллы указанных селенатов уранила).

Для изучения формальной структуры КА часто используют символические диаграммы де Брейна (*de Bruijn diagram*) [26]. Эти диаграммы описывают локальную функцию перехода КА через граф, на котором учитываются области перекрытия окрестностей соседних ячеек (рис. 7). Для $k = 3$ и $|S| = 3$ эти области будут представлять собой любые бинарные последовательности из чисел 0, 1 и 2. Очевидно, что таких последовательностей будет $|S|^{k-1} = 9$ (00, 01, 02, 10, 20, 11, 12, 21, 22) – соответствующая диаграмма де Брейна будет иметь две серии по девять вершин. Для двух перекрывающихся областей $[x - 1, x]$ и $[x, x + 1]$ (рис. 7) соответствующие им вершины диаграммы будут соединяться ребром, помеченным значением y , где

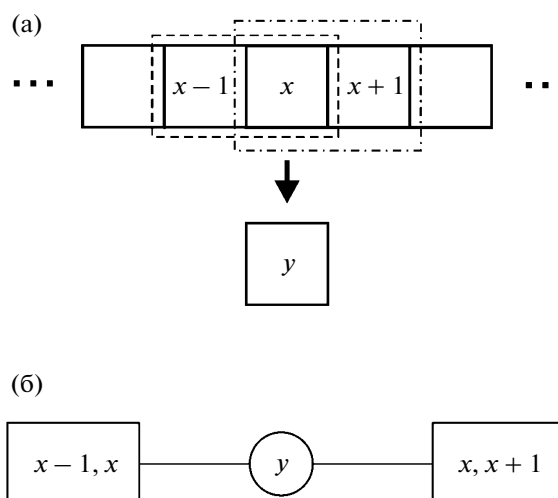


Рис. 7. К определению областей перекрытия окрестностей соседних ячеек в клеточном автомате.

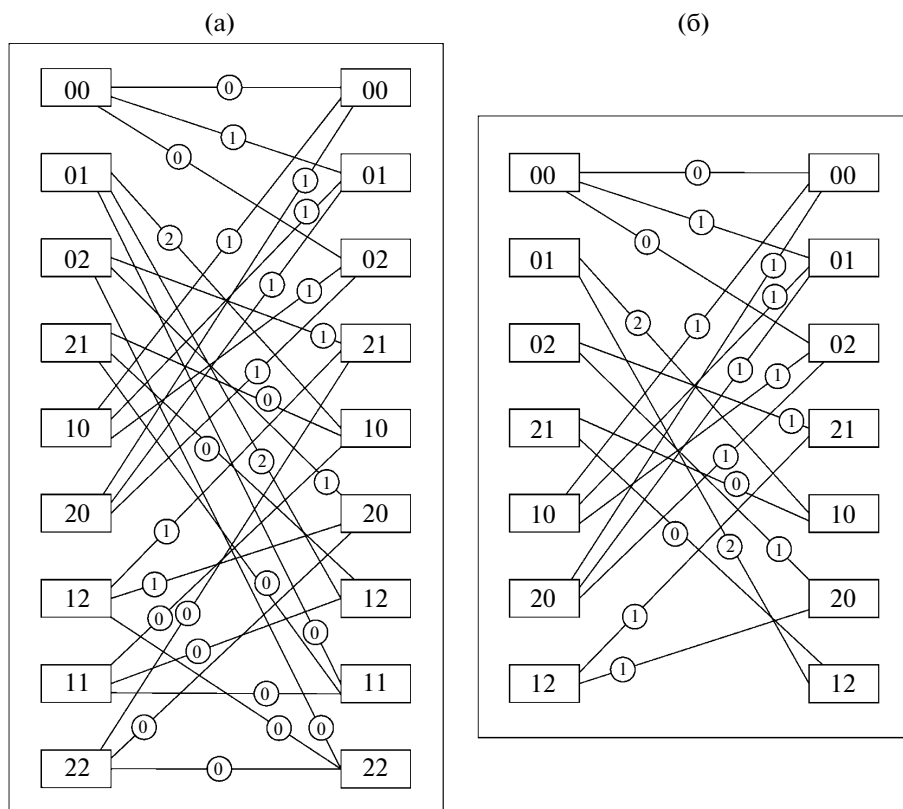


Рис. 8. Диаграмма де Брейна клеточного автомата, генерирующего ячеистые структуры на рис. 6: полный вариант (а) и вариант без орфан-конфигураций (б).

y – значение ячейки, возникающее при применении правила $[x - 1, x, x + 1] \rightarrow [y]$. Диаграмма де Брейна для КА, моделирующего рост структур селенатов уранила, изображена на рис. 8а. Исключение из нее вершин с *orphan*-конфигурациями (**11** и **22**) приводит к сокращенной диаграмме, показанной на рис. 8б. Наконец, на основе диа-

граммы де Брейна можно составить диаграмму состояния КА, показанную на рис. 9. Как можно видеть, КА состоит из семи состояний, а его алфавит представлен цифрами 0, 1 и 2. Заметим, что язык этого автомата будет воспринимать только весьма определенные строки из этого алфавита. Например, для $q_0 = 20$ допустимой строкой будет **1211010** (при этом система перейдет в состояние 10), но строка **1221010** будет недопустимой.

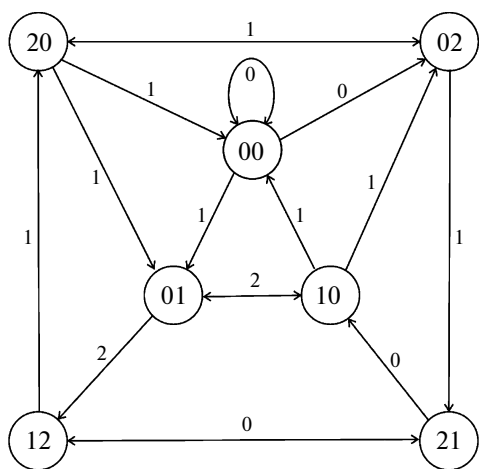


Рис. 9. Диаграмма состояния для клеточного автомата, генерирующего ячеистые структуры на рис. 6.

ЗАКЛЮЧЕНИЕ

Приложение теории автоматов к кристаллохимии и теории роста кристаллов может способствовать установлению связи между описанием структуры и описанием процесса ее роста. Использование теории конечных автоматов и их частного случая – клеточных автоматов – в структурной химии и кристаллографии позволяет установить соответствие между событиями структурообразования в физико-химических системах и абстрактными понятиями теории вычислений. При этом последовательное образование структуры можно представить как вычислительный процесс, а кристаллическую структуру – как результат выполнения определенной программы. Под этим углом зрения изучение кристаллической

структуры представляет собой изучение программы локальных взаимодействий, а кристаллограф становится специалистом в области программирования, пытающимся понять структуру чужой программы. Основную проблему современного материаловедения, таким образом, можно выразить как овладение навыками физического программирования – способностью составить и осуществить программу построения нужной физической (в частности, кристаллической) атомно-молекулярной структуры.

Работа выполнена при поддержке ФЦП “Научные и научно-педагогические кадры инновационной России” (Госконтракты 02.740.11.0326 и 16.740.11.0490).

СПИСОК ЛИТЕРАТУРЫ

1. Елисеев Э.Н., Сачков Ю.В., Белов Н.В. // Поток идей и закономерности развития естествознания. Л.: Наука, 1982. С. 171.
2. Делоне Б.Н., Долбиллин Н.П., Штогрин М.И., Галиулин Р.В. // Докл. АН СССР. 1976. Т. 227. С. 19.
3. Долбиллин Н.П. // Докл. АН СССР. 1976. Т. 230. С. 516.
4. Dolbilin N., Schattschneider D. // Field Institute Monographs. 1998. V. 10. P. 193.
5. Dolbilin N., Schattschneider D. // Field Institute Monographs. 1998. V. 10. P. 207.
6. Dolbilin N.P., Lagarias J.C., Senechal M. // Discrete Comput. Geom. 1998. V. 20. P. 477.
7. Kossel W. // Anal. Phys. 1934. B. 21. S. 457.
8. Gates D.J. // J. Chem. Soc. Faraday Trans. 1997. V. 93. P. 4223.
9. Frank F.C. // Discuss. Faraday Soc. 1949. V. 5. P. 48.
10. Brent R., Anderson M.W. // Angew. Chem. Int. Ed. 2008. Vol. 47. P. 5327.
11. Cubillas P., Anderson M.W. // Zeolites and Catalysts. Synthesis, Reactions and Applications / Eds. Cejka J. et al. Weinheim: Wiley-VCH, 2010. V. 1. P. 1.
12. Coelfen H., Antonietti M. Mesocrystals and Nonclassical Crystallization. Chichester: John Wiley & Sons, Ltd., 2008.
13. Hopcroft J.E., Motwani R., Ullman J.D. Introduction to Automata Theory, Languages and Computation. Boston: Addison-Wesley, 2001.
14. Тоффолу Т., Марголюс Н. Машины клеточных автоматов. М.: Мир, 1991. 280 с.
15. Pachinski A. Cellular Automata: A Discrete Universe. Singapore: World Scientific, 2001. 842 p.
16. Wolfram S. A New Kind of Science. Urbana: Wolfram Media Inc, 2002. 1024 p.
17. Mackay A. // Phys. Bull. 1976. V. 1976. P. 495.
18. Krivovichev S. // Acta Cryst. A. 2004. V. 60. P. 257.
19. Shevchenko V., Krivovichev S., Mackay A.L. // Физ. хим. стекла. 2010. Т. 36. С. 1.
20. Krivovichev S.V. // Eur. J. Inorg. Chem. 2010. V. 2010. P. 2594.
21. Shevchenko V., Krivovichev S. // Struct. Chem. 2008. V. 19. P. 571.
22. Krivovichev S.V., Burns P.C. // Structural Chemistry of Inorganic Actinide Compounds / Eds. Krivovichev S.V. et al. Amsterdam: Elsevier, 2008. P. 95.
23. Krivovichev S.V. // Cryst. Rev. 2004. V. 10. P. 185.
24. Krivovichev S.V. Structural Crystallography of Inorganic Oxyalts. Oxford: Oxford University Press, 2008.
25. Moore E.F. // Proc. Symp. Appl. Math. 1962. V. 14. P. 17.
26. Sutner K. // Complex Systems. 1991. V. 5. P. 19.