

УДК 530.1

SCAM: ХИМИЧЕСКИЙ КОМПЬЮТЕР

Е.И.Латкин

В в е д е н и е

Название статьи - умышленно двойственное. Слово "химический компьютер" можно толковать и как 1) специализированный процессор для моделирования химических реакций, и как 2) особая химическая реакция, имитирующая работу компьютера. Вводи - мая здесь конструкция SCAM (от Statistical Cellular Automata Machine, Машина Статистических Клеточных Автоматов) покрывает оба эти толкования.

Мы начинаем в §1 с определения модели специализированного компьютера SCAM, ориентированного на численный анализ молекулярных физико-химических процессов. Далее, в §2, мы показываем, как на SCAM можно имитировать работу любой, в том числе и универсальной машины Тьюринга. В сочетании с тезисом о физичности SCAM, выдвигаемым в §3, это приводит к высокой нижней оценке алгоритмической сложности для проблемы описания химических систем.

§1. Конструкция SCAM

За последние 2-3 десятка лет проведены успешные исследования качественного поведения химических реакций путем их имитационного моделирования методом Монте-Карло. Мы рас-

смотрим класс так называемых гетерогенных каталитических реакций, которые протекают в тончайшем слое молекул, адсорбированных на поверхности кристаллического катализатора. Для их изучения эффективны решеточные статистические модели, типа моделей Изинга, характеризующиеся предельной степенью упрощения физики явления и таким образом высвечивающие его комбинаторную сторону.

Исходя из локальных правил взаимодействия для узлов целочисленной решетки $Z \times Z$, изображающей сеть химически активных центров на поверхности катализатора, можно поставить вопрос: что можно сказать о поведении функции распределения реагентов $\xi_t: Z \times Z \rightarrow X$ в целом, в зависимости от времени $t \geq 0$, и в пределе при $t \rightarrow \infty$? Ответ на этот вопрос оказывается интересным с теоретико-алгоритмической точки зрения, и чтобы сформулировать его корректно, мы сейчас дадим строгое определение Статистического Клеточного Автомата (СКА, или, по-английски, SCA).

ОПРЕДЕЛЕНИЕ 1. *Статистическим клеточным автоматом* называется следующий набор величин:

- 1) целое число $X_0 \geq 2$ - количество возможных состояний для каждой ячейки автомата; определим $X = \{0, \dots, X_0 - 1\}$;
- 2) начальное распределение состояний $\xi_0: Z \times Z \rightarrow X$ для всех клеток автомата в момент времени $t = 0$;
- 3) таблица энергий локальных взаимодействий $E: X \times X \rightarrow R$. Если ячейка $(i_0, j_0) \in Z^2$ находится в состоянии x , а соседняя с нею ячейка (i_1, j_1) - в состоянии y , то энергия их связи равна $E(x, y)$. Соседями ячейки (i, j) считаются четыре ячейки $(i \pm 1, j)$ и $(i, j \pm 1)$. Требуется, чтобы таблица E была симметричной, т.е. $E(x, y) = E(y, x)$;
- 4) конечный набор K_1, \dots, K_M , $M \geq 1$, элементарных преобразований частичных отображений $K_m: X^2 \rightarrow X^2$. Функция K_m предписывает заменить состояния x, y пары соседних ячеек на состояния x', y' соответственно, если $K_m(x, y) = (x', y')$. Например, процес-

су диффузии веществ по поверхности катализатора отвечает функция $K_{\text{диф.}}(x, y) = (y, x)$;

5) набор вещественных чисел $k_1, \dots, k_M > 0$, показывающих скорости элементарных физико-химических процессов, описываемых преобразованиями K_1, \dots, K_M ;

6) действительное число $T > 0$ - абсолютная температура системы.

Параметры 1-6 однозначно определяют модель поведения термодинамической системы как марковский процесс; мы опишем его через алгоритм Машины SCA. Единственно, вместо температуры T нам будет удобнее пользоваться приведенной обратной температурой $\beta = 1/kT$, где k - постоянная Больцмана.

ОПРЕДЕЛЕНИЕ 2. *Машиной статистических клеточных автоматов* мы называем совокупность всех SCA вместе с алгоритмом их функционирования.

Алгоритм функционирования.

1. Вычисляется квант времени $\Delta t_0 = 1/(k_1 + \dots + k_M)$; индукцией по $s \in \mathbb{N}$ определяется распределение случайной величины $\xi_s(i, j)$ для всех $i, j \in \mathbb{Z}$ так, чтобы $\xi_s(i, j)$ соответствовала физической картине в момент времени $s \cdot \Delta t_0$. Переход $s \mapsto s+1$ называют шагом метода Монте-Карло, или МК-шагом.

2. Вводятся коэффициенты $w_m = k_m / (k_1 + \dots + k_M)$ и они считаются вероятностями для K_1, \dots, K_M . Определяется случайная величина $\mu_s^d(i, j) \in \{1, \dots, M\}$ так, что $P(\mu = m) = w_m$ и члены семейства $\{\mu_s^d(i, j)\}$ попарно статистически независимы. Символ $d \in \{+, \rightarrow, \downarrow, \uparrow\}$ указывает на одного из четырех соседей клетки (i, j) : положим $i^+ = i+1$, $i^- = i-1$, тогда $(i, j)^+ = (i^+, j)$, $(i, j)^- = (i^-, j)$, $(i, j)^{\rightarrow} = (i^+, j)$, $(i, j)^{\downarrow} = (i, j^-)$ и $(i, j)^{\uparrow} = (i, j^+)$. Числа $\mu_s^d(i, j)$ задают номер преобразования для направленной пары соседних клеток (i, j) и $(i, j)^d$ на шаге s .

3. Выбор $\pi = \mu_s^d(i, j)$ не определяет однозначно нового значения $\xi_{s+1}(i, j)$, так как K_m преобразует пару клеток, а каждая клетка (i, j) включается в четыре пары $\{(i, j), (i, j)^d\}_d$ и в восемь направленных пар вида $\langle (i, j), (i, j)^d \rangle$. Например, функция $K_m, \pi = \mu_s^d(i, j)$, может требовать замены состояния $\xi_s(i, j) = x$ на x' , а функция $K_n, \pi = \mu_s^{-d}((i, j)^d)$, - на x'' , где $-d = \leftarrow, \rightarrow$, etc. Алгоритм метода Монте-Карло устраняет эти противоречия путем поочередного перебора в среднем всех троек i, j, d в случайной последовательности.

3.1. Слова "в среднем всех" обозначают следующее. Если число ячеек конечно, как это бывает в практических вычислениях, например, $-R \leq i, j \leq R$ для некоторого $R \in \mathbb{N} \setminus \{0\}$, то конечно и число ориентированных пар соседних ячеек; оно равно $L = 2 \cdot (2R)^2$, и можно их все пронумеровать, скажем, как $\pi_R^0(0), \dots, \pi_R^0(L-1)$. Пусть $\alpha(R, l, e, s)$, где $R \geq 1, 0 \leq l < L$ и $e \in \mathbb{N}$, - набор попарно независимых случайных величин, равномерно распределенных на вещественном интервале $(0, 1)$. Тогда в конечной последовательности $v_R(1) = \lfloor \alpha(R, l, e, s) \cdot L \rfloor$, где $\lfloor \cdot \rfloor$ - целая часть числа, встречаются "в среднем все" пары соседей $\pi_R^0(v_R(1))$ из квадрата $-R \leq i, j \leq R$.

3.2. Чтобы определить последовательность перебора в бесконечной сетке $\mathbb{Z} \times \mathbb{Z}$, нам понадобятся неиспользованный параметр e и стандартная для метода Монте-Карло техника отвергнутых попыток [2]. Его суть состоит в следующем: пусть даны два размера $R_0 < R_1$ и мы перебираем пары соседей в квадрате $-R_1 \leq i, j \leq R_1$. Если выбранная пара $\pi_{R_1}^0(v_{R_1}(1))$ принадлежит квадрату $-R_0 \leq i, j \leq R_0$, мы можем объявить с вероятностью $\frac{1}{2}$, что попытка выбора не удалась, и положить номер $v_{R_1}^1(1) = -1$. Точно так же можно выбросить примерно половину пар при переборе квадрата $-R_0 \leq i, j \leq R_0$. Тогда в объединенной прореженной последова-

тельности $v_{R_0}^0(0), \dots, v_{R_0}^0(L_0 - 1), v_{R_1}^1(0), \dots, v_{R_1}^1(L_1 - 1)$, где $L_e = 2 \cdot (2R_e)^2$, каждая пара из меньшего квадрата встретится в среднем один раз, так же как и каждая пара из большего.

3.3. Организуем возрастающую последовательность размеров $R_0 < R_1 < \dots$; положим $R_e = e+1$ и договоримся, что $v(R, l, e, s) = v_{R_e}(1)$ с вероятностью $2^{-(l+e+r(R, l, e))}$ и $v(R, l, e, s) = -1$ в противном случае. Число $r(R, l, e)$ определяется как наименьшее такое R' , что $v_{R_e}(1) < 2 \cdot (2R')^2$, т.е. что пара $\pi_{R_e}^0(v_{R_e}(1))$ лежит в квадрате $-R' \leq i, j \leq R'$. Тогда подпоследовательность всех неотрицательных чисел в последовательности $\{v(R_e, l, e, s)\}_{e, l}$ перебирает в среднем все ориентированные пары ближайших соседей в бесконечной сетке ячеек $Z \times Z$. Переобозначим эту последовательность через $\{\pi_s(e)\}_{e \in \mathbb{N}}$.

4. Вводится обозначение $\xi_s^0(i, j) = \xi_s(i, j)$; случайной функцией $\pi_s(e)$ задается последовательность элементарных преобразований $\xi_s^e(i, j) \mapsto \xi_s^{e+1}(i, j)$, пределом которых будет новое распределение $\xi_{s+1}^e(i, j)$. Опишем эти преобразования. Пусть $\pi_s(e) = \langle (i, j), (i, j)^d \rangle$; число $m = \mu_s^d(i, j)$ указывает номер одной из функций K_1, \dots, K_M . Предположим, что $\xi_s^e(i, j) = x$ и $\xi_s^e((i, j)^d) = y$. Есть два обстоятельства, которые могут воспрепятствовать замене x, y на другие состояния x', y' .

4.1. Если $K_m(x, y)$ не определено, т.е. если физический процесс, описываемый функцией K_m , запрещен при таком сочетании состояний, то в этом случае никакой замены не происходит и $\xi_s^{e+1} = \xi_s^e$.

4.2. Предположим, что $K_m(x, y)$ определено и равно (x', y') . Замена состояний (x, y) на (x', y') изменяет гамильтониан системы, определяемый в статистической физике как $H[\xi] = \frac{1}{2} \sum_{i, j, d} E(\xi(i, j), \xi(i, j)^d)$. Хотя сходимость суммы для бесконеч-

ной сетки $Z \times Z$ и не очевидна, всегда определено $\Delta H_s^e = H[\zeta_s^e] - H[\xi_s^e]$, где ζ_s^e отличается от ξ_s^e измененными состояниями двух ячеек (i,j) и $(i,j)^d$. Из термодинамических соображений требуется, чтобы было либо $H_s^e < 0$, либо $\alpha_s^e < \exp(-\beta \cdot \Delta H_s^e)$, где $\{\alpha_s^e\}$ - набор независимых случайных величин, равномерно распределенных на $(0,1)$. Таким образом, если $\Delta H_s^e < 0$, то замена происходит, и $\xi_s^{e+1} = \xi_s^e$; в противном случае замена происходит с вероятностью $e^{-\beta \cdot \Delta H_s^e}$, а с вероятностью $1 - e^{-\beta \cdot \Delta H_s^e}$ замена отвергается, т.е. $\xi_s^{e+1} = \xi_s^e$.

Поскольку эволюция состояний автомата $\xi_s(i,j)$, состоящая в итерировании процедуры 4, определяет $\xi_{s+1}(i,j)$ только с точностью до статистического распределения, нам не нужна поточечная сходимость $\xi_s^e(i,j)$ при $e \rightarrow \infty$. Сходимость же распределений легко доказуема, и, таким образом, наша конструкция корректно обобщает стандартный метод Монте-Карло на бесконечную сетку $Z \times Z$.

Мы будем также рассматривать *анизотропные* статистические клеточные автоматы, подразумевая, что таблица энергий $E^d(x,y)$, а также функции преобразований $K_m^d(x,y)$ и их скорости k_m^d зависят от ориентации пары ячеек, т.е. от $d \in \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$.

ОПРЕДЕЛЕНИЕ 3. Машина анизотропных СА отличается формулой для $\Delta t_0 = \sum_{m,d} k_m^d$ и $w_m^d = k_m^d / \Delta t_0$.

Говоря о конкретном автомате, мы будем иногда опускать характеристику "изотропный" или "анизотропный", поскольку его тип очевиден из контекста.

§2. САМ и машины Тьюринга

В статье [3], посвященной анализу сложности предельных языков некоторых клеточных автоматов (СА = Cellular Automata), в частности, показано, как средствами СА можно вычислить пове -

дение любой машины Тьюринга. Отсюда следует, что вычислимость на CA , по крайней мере, не слабее рекурсивной вычислимости. Мы намерены доказать то же самое для статистических клеточных автоматов: работу любой машины Тьюринга можно имитировать в рамках $SCAM$.

Воспользуемся конструкцией из [3], которая опирается на теорему из [6] о том, что рекурсивная вычислимость эквивалентна вычислимости на так называемых *программируемых машинах* с двумя регистрами A и B , способными запоминать по одному натуральному числу.

Программируемая машина состоит из конечной последовательности инструкций трех типов:

- 1) R^+ - прибавить единицу к регистру $R \in \{A, B\}$;
- 2) $R^-[n]$ - отнять единицу от R , если $R > 0$; иначе - перейти к инструкции номер n ;
- 3) $[n]$ - безусловный переход к n -й инструкции.

Инструкции нумеруются от 1 до N ; команда перейти к инструкции номер $n = 0$ или $n > N$ интерпретируется как конец вычислений. Параметры вычисляемой функции помещают в регистры перед запуском машины, а после ее остановки - считывают результат.

В [3] строится 1-мерный клеточный автомат (Ξ, F) , где $\Xi: (i, s) \in \mathbb{Z} \times \mathbb{N} \mapsto \Xi_s(i) \in X$ - состояние i -й ячейки на шаге s , и $(F(\Xi_s))_i = \Xi_{s+1}(i)$ - эволюционная функция, причем локальная, т.е. состояние $\Xi_{s+1}(i) = f(\Xi_s(i-r), \dots, \Xi_s(i), \dots, \Xi_s(i+r))$ зависит только от $2r+1$ ближайших ячеек.

Автомат (Ξ, F) , имитирующий произвольно выбранную программируемую машину M_0 , выглядит следующим образом. В начальном состоянии ячейки $\Xi_0(0)$ закодирована первая команда машины M_0 ; слева от нее записывается n_A символов "A", где n_A - начальное значение регистра A , и, далее, для $i \leq -(n_A+1)$ все $\Xi_0(i) = "a"$;

аналогично, справа от центра загружается начальное состояние регистра В символами "В" и "b".

Мы несколько отклонились от конструкции [3] в нужную для нас сторону, но основная идея остается та же. Можно так запрограммировать таблицу переходов $f(\dots, \Xi_s(i), \dots)$, чтобы состояние C_n в ячейке $i = 0$, соответствующее n -й команде машины M_0 , управляло нужными действиями над количеством символов "А" и "В", записанных в массив Ξ . Пусть, например, $C_n = B^-[m]$ - команда вычест единицу из регистра В; ее можно выполнить, стерев самую правую большую букву "В" в Ξ и заменив ее на маленькую "b". Для этого на шаге $s \rightarrow s+1$ мы повесим метку "минус" на символ "В" в ячейке $i = 1$, т.е. заменим его новым символом "В⁻". (Если $\Xi_s(i=1) = "b"$, то C_n сразу переходит в новое состояние C_m .)

Состояние "В⁻" распространяется дальше вправо, пока не достигнет ячейки, занятой "b". Здесь последний символ "В" съедается - заменяется на "b", а влево отправляется сигнал "В*", который, достигнув ячейки $i = 0$, заставит ее перейти в состояние C_{n+1} или в состояние "Halt", которое зарезервировано на случай, если $n = N$ - последняя команда в программе M_0 .

Очевидно описание этой процедуры в рамках одномерной изотропной SCAM. Достаточно завести лишь одно правило $K(x, y) \mapsto (x', y')$, преобразующее пары:

- | | |
|------------------------------------|--|
| 1. $(C_n, b) \mapsto (C_m, b)$; | 5. $(B^-, b^*) \mapsto (B^*, b)$; |
| 2. $(C_n, B) \mapsto (C_n, B^-)$; | 6. $(B^-, B^*) \mapsto (B^*, B)$; |
| 3. $(B^-, B) \mapsto (B^-, B^-)$; | 7. $(C_n, B^*) \mapsto (C_{n+1}, B)$; |
| 4. $(B^-, b) \mapsto (b^*, b)$; | 8. $(C_n, b^*) \mapsto (C_{n+1}, b)$. |

Нетрудно продолжить таблицу и для других типов команд: $B^+, A^+, A^-[m]$ и $[m]$. Физическая скорость k процесса $K(x, y)$ и энергии связей $E(x, y)$ не играют роли.

Таким образом, доказано следующее утверждение.

ТЕОРЕМА о вложении. SCAM может имитировать работу универсальной машины Тьюринга.

Мы, однако, хотим большего: проинтерпретировать работу M_0 на такой SCAM, которую можно было бы признать физически реализуемой. Понимается, никакая химическая система не может иметь бесконечной пространственной протяженности, но мы собираемся проследить за соблюдением двух важнейших законов: сохранения вещества и сохранения энергии.

ОПРЕДЕЛЕНИЕ 4. Статистический клеточный автомат называется обратимым, если соблюдается хотя бы относительный баланс прямых и обратных элементарных преобразований. Обозначим через $M(x, y; x', y')$ множество $\{m | K_m(x, y) = (x', y')\}$, где K_1, \dots, K_M - набор локальных трансформаций для автомата, и для всех x, y и x', y' положим $W(x, y; x', y') = \sum_m w_m$, где сумма берется по всем $m \in M(x, y; x', y')$. Тогда из $W(x, y; x', y') > 0$ должно следовать, что вероятность обратного превращения $W(x', y'; x, y) > 0$. Класс обратимых SCA мы обозначим через $SCAM_0$.

Вопрос о соотношении вычислимости на $SCAM_0$ и рекурсивной вычислимости, к сожалению, слишком объемный для данной небольшой статьи, носящей анонсирующий характер. Отметим лишь те основные трудности, которые мы намерены преодолеть в последующих публикациях, чтобы показать равносильность машин Тьюринга и $SCAM_0$.

С одной стороны, если рассмотреть распространение метки "B⁻" в нашем примере, считая, что каждый шаг в нем реализуется переходом молекулы "B" в другое квантовое состояние "B⁻" или присоединением к ней какого-то атома, придется признать, что это движение не может длиться неограниченно долго, так как оно связано либо с поглощением, либо с выделением энергии и/или вещества. Требование, чтобы автомат был локально обра-

тим, запрещает направленное распространение сигналов по этим физическим соображениям, оставляя возможной лишь их диффузию.

Если бы требование обратимости не относилось к ячейке $i = 0$, то диффузии сигналов вполне хватило бы, чтобы организовать имитацию M_0 на $SCAM_0$: сигнал "минус", блуждая туда-сюда в интервале, зажатом между ячейками $i = 1$ и $i = p_B + 1$, рано или поздно достигает последней, причем в обозримое в среднем время. От диффузионной связи пострадает лишь скорость работы машины, но не ее выразительная сила.

С другой стороны, если не потребовать обратимости, то поведение $SCAM$ будет заведомо рекурсивно невычислимо. В самом деле, легко представить себе статистический клеточный автомат, все ячейки которого при $t = 0$ находятся в состоянии " \perp " = "неопределенность". На первом же шаге состояние " \perp " имеет право спонтанно превратиться либо в 0, либо в 1, и таким образом может возникнуть бесконечное битовое поле сколь угодно высокой степени сложности. Условие локальной обратимости, хоть это и не так очевидно, способно послужить лекарством от подобных неприятностей.

§3. Тезис о физичности $SCAM$

Причина, заставляющая нас уделять особое внимание именно обратимым статистическим клеточным автоматам, состоит в том, что они могут претендовать на физическую реалистичность. Это утверждение, которое мы хотели бы выдвинуть в качестве тезиса, очень важно, так как позволяет увязать теоретико-алгоритмические рассуждения в классе $SCAM_0$ с актуальными проблемами теоретической химии.

ТЕЗИС о физичности $SCAM$. Конструкция обратимых статистических клеточных автоматов не противоречит законам физики.

Опишем выкладки, позволяющие интерпретировать автоматы класса SCAM₀ как макроканонические ансамбли Гиббса [4], а алгоритм их работы - как процедуру Метрополиса для поиска их равновесной конфигурации [1].

В самом деле, чтобы алгоритм SCAM совпадал с алгоритмом Метрополиса, достаточно, чтобы соблюдалось условие детального равновесия: $e^{\beta \cdot \Delta H} \cdot W_e(x, y; x', y') = W_e(x', y'; x, y)$, где через $W_e(x, y; x', y')$ обозначено $\sum_m e^{\beta \cdot N(m)} \mu(m)$ - сумма по всем $m \in M(x, y; x', y')$. Здесь ΔH - скачок суммарной энергии связей при замене x, y на x', y' и $N(m), \mu(m)$ - соответственно концентрация и химический потенциал некоего квазивещества "m", с которым можно увязать процесс K_m .

Коэффициенты $N(m)$ и $\mu(m)$ произвольны в том смысле, что любой их набор можно считать соответствующим физической реальности. Обратимость автомата необходима, чтобы можно было, подобрав их, привести автомат в равновесие. С другой стороны, занимаясь теорией вычислимости на SCAM, мы располагаем достаточной свободой в варьировании скоростей элементарных процессов, энергий связи и этих дополнительных параметров, чтобы уравновесить практически любой интересующий нас обратимый автомат.

§4. Молекулярная синергетика

В заключение отметим, что возникновение необратимого во времени поведения у большой системы, состоящей из обратимых элементов, часто встречающаяся и важная закономерность в физике и химии [7]. Строго говоря, локально обратим любой электрохимический процесс, так как он подчиняется обратимым по переменной t уравнениям Шредингера и Максвелла. Вместе с тем, работа мозга доставляет пример довольно сложного и в целом необратимого электрохимического процесса.

Изучение сложных систем, состоящих из множества простых элементов, составляет предмет синергетики [5], а применительное к молекулярной химии ее направление можно обозначить как *молекулярную синергетику*. Смысл нашей статьи состоит в попытке теоретико-алгоритмического рассмотрения данной проблематики.

Основное утверждение, которое мы намерены уточнить в наших ближайших публикациях, состоит в следующем: в принципе возможна такая химическая система, которая могла бы имитировать работу универсальной машины Тьюринга.

Следствие: поведение химических систем, вообще говоря, нельзя описать аналитически (в элементарных функциях). По сути, теория химической кинетики должна быть эквивалентна теории алгоритмов.

Л и т е р а т у р а

1. БИНДЕР К. (ред.). Методы Монте-Карло в статистической физике. - М.: Мир, 1982.
2. ЕРМАКОВ С.М., МИХАЙЛОВ Г.А. Статистическое моделирование. - М.: Наука, 1982. - 295 с.
3. GOLES E., MAASS A., MARTINEZ S. On the limit set of some universal cellular automata // Theoretical Computer Science. - 1993. - Vol. 110. - P. 53-78.
4. КИТТЕЛЬ Ч. Статистическая термодинамика. - М.: Наука, 1977. - 336 с.
5. ЛОСКУТОВ А.Ю., МИХАЙЛОВ А.С. Введение в синергетику. - М.: Наука, 1990. - 272 с.
6. MINSKY M. Computation: Finite and Infinite Machines. - Prentice-Hall (Englewood Cliffs, NJ), 1967.
7. ПРИГОЖИН И., СТЕНГЕРС И. Порядок из хаоса. - М.: Прогресс, 1986. - 432 с.

Поступила в редакцию
30 января 1995 г.