

УДК 519.115:519.2

АНАЛИЗ СХЕМЫ РАЗМЕЩЕНИЯ НЕРАЗЛИЧИМЫХ ЧАСТИЦ ПО НЕРАЗЛИЧИМЫМ ЯЧЕЙКАМ

Н. Ю. Энатская, Е. Р. Хакимуллин, А. В. Колчин

*Московский институт электроники и математики
Национального исследовательского университета «Высшая школа экономики»*

В общей схеме размещения неразличимых частиц по неразличимым ячейкам и в частном случае схемы без пустых ячеек проводятся следующие исследования. Для общего числа исходов частной схемы выписано рекуррентное соотношение, и по нему получена явная точная формула для него. Выявлена связь чисел исходов в общей и частной схемах. Описан случайный процесс последовательного размещения по одной частице по ячейкам, на основании которого приводится алгоритм решения задачи перечислительной комбинаторики представления всех возможных исходов размещения фиксированного числа частиц по ячейкам в схеме и получения распределения их вероятностей. Предложены разные способы моделирования состояний схемы и приближенного нахождения числа ее исходов методом стохастического моделирования.

Ключевые слова: перечислительные задачи комбинаторного анализа, схема размещения.

N. Yu. Enatskaya, E. R. Khakimullin, A. V. Kolchin. ANALYSIS OF A SCHEME OF ALLOCATING INDISTINGUISHABLE PARTICLES TO INDISTINGUISHABLE CELLS

The following investigations are carried out in the general scheme of allocation of indistinguishable particles to indistinguishable cells and in the particular scheme with no empty cell allowed. A recurrence relation is found for the total number of outcomes of the particular scheme and an explicit expression for it is obtained. A relation between the numbers of outcomes of the general and particular schemes is found. A random process of successive allocation of a single particle to cells which provides us with an algorithm for solving the combinatorial problem and finding all the outcomes of allocation of a fixed number of particles to cells in the scheme under consideration and finding the distribution of their probabilities is described. Various methods to simulate the states of the scheme and approximate the number of its outcomes by means of stochastic simulation techniques are suggested.

Key words: enumerative combinatorics, allocation scheme.

1. ВВЕДЕНИЕ

Комбинаторные задачи и методы занимают значительное место в исследованиях по теории вероятностей. В последние десятилетия замет-

ные усилия были направлены на изучение задач о случайном размещении частиц по ячейкам, где r дробинки независимо друг от друга бросают в n ящиков. Внимание к задачам о

размещении частиц является отражением общего повышения интереса к комбинаторным задачам, который вызван в том числе возросшими перечислительными возможностями вычислительных машин. Задачи изучения случайных размещений частиц по ячейкам возникают в статистической физике, вычислительной технике, теории автоматов, в математической физике, астрономии и биологии, и не раз независимо переоткрывались многими авторами в различных теоретических и прикладных исследованиях.

1.1. Вид результата размещения частиц по ячейкам

Пусть n — число ячеек в рассматриваемой схеме. Результаты размещения в схеме размещения неразличимых частиц по неразличимым ячейкам описываются векторами вида $\bar{\mu}_r = (\mu_0, \mu_1, \dots, \mu_r)$, где r — число частиц, μ_i — число ячеек, содержащих ровно по i частиц после размещения r частиц по n ячейкам, так что справедливы соотношения

$$\sum_{i=0}^r \mu_i = n, \quad \sum_{i=1}^r i\mu_i = r.$$

Результат размещения можно также представить в виде вариационного ряда заполнения ячеек вида $(0, \dots, 0, 1, \dots, 1, \dots, r)$, где число значений i есть μ_i , $i = 0, \dots, r$. В частном случае схемы без пустых ячеек $\mu_0 = 0$.

1.2. Сравнение схем размещения неразличимых частиц по различным и неразличимым ячейкам

В нашей схеме размещения с неразличимыми ячейками исходы описываются векторами $\bar{\mu}_r$, в то время как в схеме размещения с различными ячейками они определяются различными заполнениями ячеек с учетом их порядка. Поэтому число исходов $N^* = N^*(r, n)$ в нашей схеме меньше, чем число исходов $N_1 = N_1(r, n)$ в схеме размещения с различными ячейками за счет того, что каждому фиксированному исходу, описываемому вектором $\bar{\mu}_r$ в рассматриваемой нами схеме, соответствует $n! / \prod_{i=0}^r \mu_i!$ исходов в схеме размещения с различными ячейками.

1.3. Изменение вектора $\bar{\mu}_k$ при добавлении одной частицы

Опишем отличие вектора $\bar{\mu}_{k+1}$ от вектора $\bar{\mu}_k$ при размещении по n ячейкам после k еще одной частицы. Оказывается, что при этом в векторе $\bar{\mu}_k$ будут изменяться на единицу ровно две соседних компоненты, причем левая будет уменьшаться на единицу, а правая — увеличиваться на единицу. В силу неотрицательности

компонент векторов $\bar{\mu}_k$ и $\bar{\mu}_{k+1}$ число разных векторов $\bar{\mu}_{k+1}$, получающихся из данного вектора $\bar{\mu}_k$, будет равняться числу его ненулевых компонент, причем это число в векторе $\bar{\mu}_{k+1}$, по сравнению с вектором $\bar{\mu}_k$, при попадании добавленной частицы в ячейку с i частицами, в зависимости от значений компонент вектора $\bar{\mu}_k$,

- а. не изменяется при $\mu_i > 1, \mu_{i+1} > 0$;
- б. увеличивается на 1 при $\mu_i > 1, \mu_{i+1} = 0$;
- в. уменьшается на 1 при $\mu_i = 1, \mu_{i+1} > 0$.

Нетрудно видеть, что из разных векторов $\bar{\mu}_k$ при добавлении одной частицы могут получаться одинаковые векторы $\bar{\mu}_{k+1}$. Приведем соответствующий пример.

Пример 1. Пусть $n = 4$, тогда $\bar{\mu}_5 = (0, 3, 1, 0, 0, 0)$, и $\bar{\mu}_5 = (1, 2, 0, 1, 0, 0)$ приводят к вектору $\bar{\mu}_6 = (0, 3, 0, 1, 0, 0, 0)$ при попадании частицы в первом случае в ячейку с двумя частицами, а во втором — в пустую.

2. Точное вычисление чисел исходов в общей и частной схемах размещения частиц по ячейкам

2.1. Связь чисел исходов в общей и частной схемах размещения

Число исходов $N^* = N^*(r, n)$ в общей схеме может быть получено из числа $N = N(r, n)$ частной схемы без пустых ячеек. Для этого предлагается следующая процедура: к данным r частицам добавим n частиц, тем самым получив $r^* = r + n$ частиц, которые разместим по n ячейкам без пустых ячеек $N(r^*, n) = N(r+n, n)$ способами. После этого извлечем из каждой ячейки по одной частице, отчего число размещений не изменится и снова будет равно $N(r+n, n)$. Отсюда получаем, что

$$N^*(r, n) = N(r+n, n). \quad (1)$$

2.2. Рекуррентная формула для $N = N(r, n)$

Для $N = N(r, n)$ найдем рекуррентное соотношение с помощью деления всей совокупности N исходов частной схемы на две, численности которых зависят от меньших значений параметров, чем исходные r и n . К первой совокупности отнесем те исходы из $N(r, n)$, в которых минимальное заполнение ячеек равно 1; их число равно $N(r-1, n-1)$, и они получаются, если в любую ячейку положить одну частицу, а остальные $r-1$ частиц разместить по той же схеме по остальным $n-1$

ячейкам. Ко второй отнесем те исходы, в которых минимальное заполнение ячеек строго превосходит 1; их число равно $N(r - n, n)$, и они получаются, если во все ячейки положить по одной частице, а остальные $r - n$ частиц разместить по n ячейкам по той же схеме. Таким образом, приходим к рекуррентной формуле

$$N(r, n) = N(r - 1, n) + N(r - n, n), \quad (2)$$

где числа $N(r, n)$ обладают при $r, n \geq 0$ следующими очевидными свойствами:

1. $N(r, r) = 1$;
2. $N(r, n) = 0$ при $r < n$;
3. $N(r, 0) = 0$;
4. $N(r, 1) = 1$.

Используя приведенные выше свойства чисел $N(r, n)$, называемые далее предельными, проводим по (2) такое число итераций по каждой ветви рекурренты (2), которое приводит к предельным значениям слагаемых. При этом, в общем случае, длины ветвей рекурренты до достижения предельных значений по свойствам 1–4 различны.

Пример 2. Пусть $r = 7, n = 3$. Тогда по (2) получаем, что

$$\begin{aligned} N(7, 3) &= N(6, 2) + N(4, 3) = \\ &= N(5, 1) + N(4, 2) + \\ &\quad + N(3, 2) + N(1, 3) = \\ &= 1 + N(3, 1) + N(2, 2) + N(2, 1) + \\ &\quad + N(1, 2) + 0 = \\ &= 1 + 1 + 1 + 1 + 0 + 0 = 4. \end{aligned}$$

Элементарный перебор исходов схемы дает следующие различные заполнения ячеек, без пустых, с учетом неразличимости ячеек: $(1, 1, 5)$, $(1, 2, 4)$, $(1, 3, 3)$, $(2, 2, 3)$, что соответствует векторам

$$\begin{aligned} \bar{\mu}_7 &= (\mu_0, \mu_1, \mu_2, \mu_4, \mu_5, \mu_6, \mu_7) \\ &\in \{(0, 2, 0, 0, 0, 1, 0, 0), \\ &\quad (0, 1, 1, 0, 1, 0, 0, 0), \\ &\quad (0, 1, 2, 0, 0, 0, 0, 0), \\ &\quad (0, 0, 2, 1, 0, 0, 0, 0)\}, \end{aligned}$$

иными словами, получаем четыре варианта, как и по рекуррентной формуле (2).

2.3. Точная формула для $N = N(r, n)$

Вывод явной формулы для $N(r, n)$ проведем на основе рекуррентного соотношения (2) и закономерностей его применения при многократных итерациях с понижением значений параметров r и n до получения предельных значений.

Замечание 1. При получении явной формулы для числа $N(r, n)$ будем считать, что его предельные значения по параметрам достигаются на одной и той же m -й итерации по всем ветвям рекурренты (2), несмотря на то что при численном ее применении для расчета числа $N(r, n)$ это не всегда так (см. пример 2), поскольку ветви рекурренты могут заканчиваться по свойствам 1–4 числа N . Тем не менее наше допущение выполняется, если формально применять рекуррентную формулу (2) к числам $N(r, n)$, уже достигшим предельных значений, для чего доопределим свойства чисел $N(r, n)$ для параметров r и n , допускающих значения, меньшие или равные нулю, следующим образом:

- 1а. $N(r, r) = 1, r \geq 0$;
- 2а. $N(r, n) = 0, r < n$;
- 3а. $N(r, n) = 0, n \leq 0, r + n \neq 0$;
- 4а. $N(r, 1) = 1$.

Это значит, что количества итераций по всем ветвям рекурренты (2) могут быть взяты равными.

Введем индикаторную функцию

$$I(z) = \begin{cases} 0, & z < 0; \\ 1, & z \geq 0. \end{cases}$$

Тогда искомое число N , с учетом свойств 1а–4а и рекуррентной формулы (2), может быть представлено в виде

$$N(r, n) = \sum_{j=1}^{2^m} I(r_{mj} - n_{mj}), \quad (3)$$

где

$$m = \min \left\{ i : \sum_{j=1}^{2^i} (n_{ij} - 1)(r_{ij} - n_{ij})I(r_{ij} - n_{ij}) = 0 \right\}, \quad (4)$$

и где r_{ij} и n_{ij} есть соответственно значения параметров r и n в j -м слагаемом по рекуррентной формуле (2) при i -й итерации. Для единства обозначений положим

$$r = r_{01}, \quad n = n_{01}.$$

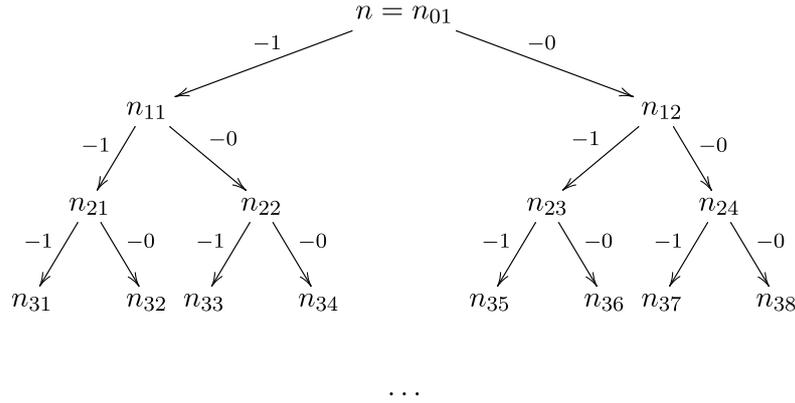


Рис. 1. Схема 1

Формула (4) дает способ нахождения числа m путем проверки для каждого $i = 1, 2, \dots$ условия

$$\sum_{j=1}^{2^i} (n_{ij} - 1)(r_{ij} - n_{ij})I(r_{ij} - n_{ij}) = 0, \quad (5)$$

при первом выполнении которого получаем число m , равное последнему значению i .

Таким образом, задача сводится к построению явных формул для вычисления параметров r_{mj} и n_{mj} , участвующих в формуле (3), для чего требуется находить значения r_{ij} и n_{ij} и для $i < m$, причем с учетом характера ветвления рекурренты (2) на два слагаемых; на i -й итерации индекс j меняется от 1 до 2^i .

Для выявления закономерностей при применении рекуррентной формулы (2) приведем схемы изменения каждого из параметров r и n в слагаемых при последовательных итерациях рекурренты (2). Начнем с более простой схемы 1 (рис. 1) для параметра n_{ij} . Для этого по вертикали (сверху вниз) представим параметры n_{ij} по мере проведения итераций, а по горизонтали — его последовательные обозначения в слагаемых рекуррентного соотношения в порядке его применения. На стрелках схемы перехода к следующей итерации укажем действия для пересчета параметра n_{ij} .

Будем рассматривать схему 1 как граф переходов из состояния n_{01} во все другие возможные по рекурренте (2) состояния n_{ij} . Всем правым переходам (ребрам) в графе поставим в соответствие нули, а левым — единицы. Тогда переход по (2) из состояния n_{01} в любое возможное состояние на m -м уровне (то есть на m -й итерации) можно представить траекторией $(T_m) = (T_m)_2$ в виде последовательности нулей и единиц в соответствии с числами

на последовательных ребрах траектории от состояния n_{01} до состояния n_{mj} , а полученная последовательность (T_m) будет представлять собой двоичную запись числа $2^m - j$, где j — номер состояния на m -м уровне. По схеме 1 пересчета значения n_{mj} из $n = n_{01}$ с учетом операций, указанных на схеме по итерациям рекурренты, следует, что для вычисления значения n_{mj} из начального значения n нужно вычесть столько единиц, сколько их имеется в траектории графа от состояния $n = n_{01}$ до состояния n_{mj} , то есть число единичных разрядов двоичной записи числа $2^m - j$. Отсюда получаем формулу для вычисления n_{mj} :

$$n_{mj} = n - S_{2^m - j}, \quad (6)$$

где S_z — сумма двоичных разрядов числа z .

Покажем применение формулы (6) на примере.

Пример 3. При $m = 3$, состояниям с номерами от 1 до 2^3 , то есть от 1 до 8, определяемым значениями параметра n_{3j} , то есть состояниям $n_{31}, n_{32}, n_{33}, n_{34}, n_{35}, n_{36}, n_{37}, n_{38}$, будут соответствовать траектории (111), (110), (101), (100), (011), (010), (001), (000), представляющие собой двоичные записи чисел $(2^m - j)$, где j — номер состояния на уровне $m = 3$. Тогда $\{n_{3j}\}$ при $j = 1, \dots, 8$ будут принимать в порядке роста индекса j значения $\{n - 3, n - 2, n - 2, n - 1, n - 2, n - 1, n - 1, n\}$, что совпадает с результатом непосредственных вычислений по схеме 1 по указанным операциям на ребрах графа по каждой траектории.

Замечание 2. Формула (4) при замене числа m на любое i , не превосходящее m , дает возможность определять значение n_{ij} на i -й итерации.

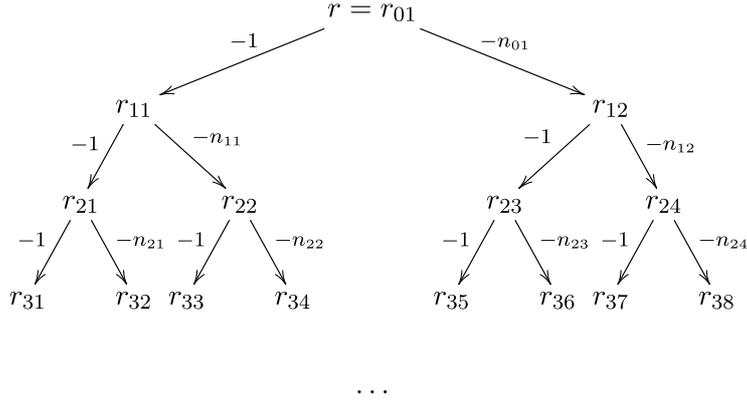


Рис. 2. Схема 2

Замечание 3. Обозначив через $(T_i) = (T_i)_2$ запись числа T_i в двоичном коде и через $(T_i)_{10}$ запись числа в десятичном коде, соответствующие траектории (T_i) до i -го уровня включительно, получаем, что номер слагаемого на i -й итерации, то есть второй индекс последнего состояния траектории (T_i) на i -м уровне, удовлетворяет равенству

$$j = 2^i - (T_i)_{10}. \quad (7)$$

Поясним использование результата (7) на примере.

Пример 4. Пусть $i = 3$ (см. схему 1). Тогда траектория (111) пересекает уровень $i = 3$ по состоянию n_{3j} , где в силу (7)

$$j = 2^3 - (111)_{10} = 8 - 7 = 1;$$

траектория (110) пересекает уровень $i = 3$ по состоянию n_{3j} , где в силу (7)

$$j = 2^3 - (110)_{10} = 8 - 6 = 2;$$

траектория (101) пересекает уровень $i = 3$ по состоянию n_{3j} , где в силу (7)

$$j = 2^3 - (101)_{10} = 8 - 5 = 3,$$

и так далее.

Аналогично рассмотрим изменения первого параметра числа $N(r, n)$, именно – параметра r_{ij} при последовательных итерациях рекурренты (2). Приведем соответствующую схему (аналогичную схеме 1) пересчета значений r_{ij} по итерациям рекурренты (2) с указанием операций на стрелках переходов (ребрах графа).

Снова, как и схему 1, рассматриваем схему 2 (рис. 2) как граф переходов из состояния $r = r_{01}$ во все другие возможные состояния при итерациях по рекурренте (2) и будем ставить в соответствие всем правым переходам (ребрам графа) нули, а левым – единицы. Тогда переход из состояния r_{01} в любое j -е состояние на m -м уровне r_{mj} можно представить траекторией $(T_m) = (T_m)_2$ – последовательностью нулей и единиц в порядке итераций, то есть двоичной записью числа $2^m - j$. Из схемы 2 следует, что для вычисления значения r_{mj} из начального значения $r = r_{01}$ нужно вычесть столько единиц, сколько их имеется в траектории (T_m) от r_{01} до r_{mj} , то есть число S_{2^m-j} и число вида $\sum_{k=1}^{L_m} n_{p_k-1, l_k}$, где L_m – число нулей в траектории (T_m) , $\{p_k\}$ – места нулей в траектории (T_m) , а значение l_k вычисляется по формуле (7) как точка пересечения траектории (T_m) с уровнем $p_k - 1$. То есть по (7) получаем

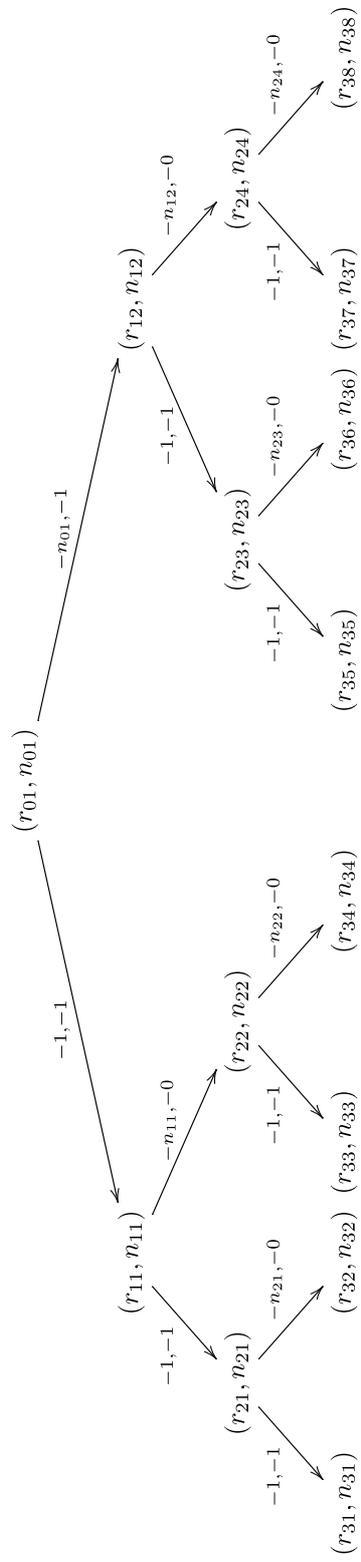
$$l_k = 2^{p_k-1} - (T_{p_k-1})_{10},$$

где $(T_i)_{10}$ – десятичный код числа (T_i) . Отсюда получаем, что

$$r_{mj} = r - S_{2^m-j} - \sum_{k=1}^{L_m} n_{p_k-1, l_k}, \quad (8)$$

где значения n_{p_k-1, l_k} вычисляются по формуле (6).

Замечание 4. Формула (8) дает общий вид выражения для r_{ij} при замене в ней m на i , а L_m – на L_i .



...

Рис. 3. Схема 3

Формулы (3)–(8) дают явное точное выражение для определения искомого числа исходов $N(r, n)$ в частной схеме размещения r неразличимых частиц по n неразличимым ячейкам без пустых ячеек. Тогда (см. п.2.1) число исходов $N^*(r, n)$ в общей схеме, допускающей пустые ячейки, пересчитывается из числа $N(r, n)$ по формуле (1).

Поясним порядок использования формул (3)–(8) при вычислении числа $N(r, n)$ по формуле (3).

Для наглядности расчета приведем общую схему пересчета параметров r_{ij} и n_{ij} по итерациям рекурренты (2). Рассматривая ее как граф переходов из состояния $r = r_{01}, n = n_{01}$ во все другие возможные по рекурренте (2) состояния, будем указывать на ребрах графа при левых и правых переходах операции пересчета параметров r и n в порядке этих переходов.

Вычисление искомого числа $N = N(r, n)$ начинаем с определения числа m — номера итерации, когда числа N впервые по всем ветвям итераций рекурренты достигают предельных значений по обобщенным свойствам 1а–4а чисел $N(r, n)$.

Оказывается, что вместо трудоемкого глобального вычисления параметров r_{ij} и n_{ij} по схеме 3 (рис. 3) для проверки выполнения условия (5), при выполнении которого впервые последнее значение индекса i фиксируется в качестве искомого числа m , можно определить значение m сразу непосредственно по данным r и n . Для этого введем понятие длины траектории в графе без выравнивания длин ветвей рекурренты, понимаемое как число итераций до достижения параметрами r и n значений, при которых число N принимает предельное значение по свойствам 1–4.

Длина траектории определяется значениями параметров r и n и скоростями их убывания по итерациям до выполнения впервые хотя бы одного из условий: $n = 1$ или $r \leq n$.

В качестве искомого числа m нас интересует длина наибольшей траектории. В соответствии с рекуррентной формулой (2), для числа $N(r, n)$ по левым сдвигам оба параметра убывают на любой итерации на единицу, а по правым — убывает только первый параметр r на величину второго параметра, то есть тем медленнее, чем меньше второй параметр n . С учетом очевидного неравенства

$$\frac{r}{n} < \frac{r-k}{n-k}, \quad k > 0,$$

при $r > n$ это означает, что траектория максимальной длины должна состоять вначале из максимального числа $n - 2$ левых переходов, не приводящих ее к концу (при $(n - 1)$ -м левом

переходе (сдвиге) параметр n принимает значение 1, и траектория заканчивается), чтобы максимально уменьшить скорость убывания по итерациям параметра r , а затем из правых переходов, число которых и будем определять. К началу правых переходов первый параметр принимает значение $r - (n - 2) = r - n + 2$ и убывает при каждом правом сдвиге на величину второго параметра, который достигает значения 2 и при правых переходах не меняется. Траектория по подряд идущим правым сдвигам заканчивается в соответствии со свойствами 1–4 чисел $N(r, n)$, когда первый параметр впервые становится меньшим или равным двум. Значит, число правых переходов есть $[(r - n + 2)/2]$, если число $r - n + 2$ четно, или $[(r - n + 2)/2] + 1$ в противном случае, что, очевидно, записывается эквивалентной формулой $[(r - n + 3)/2] - 1$, где $[z]$ есть целая часть числа z . Тогда искомая максимальная длина m вычисляется из исходных значений r и n по формуле

$$\begin{aligned} m &= n - 2 + \left[\frac{r - (n - 2) + 1}{2} \right] - 1 = \\ &= n - 3 + \left[\frac{r - n + 3}{2} \right]. \end{aligned} \quad (9)$$

Легко проверяется, что любое изменение порядков и количеств левых сдвигов может только уменьшить длину траектории.

Далее после нахождения числа m выписываются подряд все траектории $(T_m)_2$, от первой до 2^m -й, в виде последовательностей нулей, соответствующих правым сдвигам, и единиц, соответствующих левым сдвигам. По каждой j -й траектории, $j = 1, \dots, 2^m$, находятся число нулей L_m и их места $\{p_k\}$, $k = 1, \dots, m$, в траектории, и по формулам (6) и (8) вычисляются значения параметров n_{mj} и r_{mj} , после чего для нахождения искомого числа $N = N(r, n)$ используется формула (3) с учетом свойств 1а–4а.

Проведем вычисление числа N по описанному алгоритму на примере.

Пример 5. Пусть $r = 7, n = 3$. (Результат примера 5 был получен в примере 2 непосредственным расчетом по рекуррентной формуле (2) и свойствам 1–4 чисел $N(r, n)$). Здесь вначале производим непосредственное вычисление числа N с выравниванием длин траекторий до m по свойствам 1а–4а по схемам 1, 2, 3 и по рекуррентной формуле (2). По формуле (9)

$$m = 3 - 3 + \left[\frac{7 - 3 + 3}{2} \right] = 3.$$

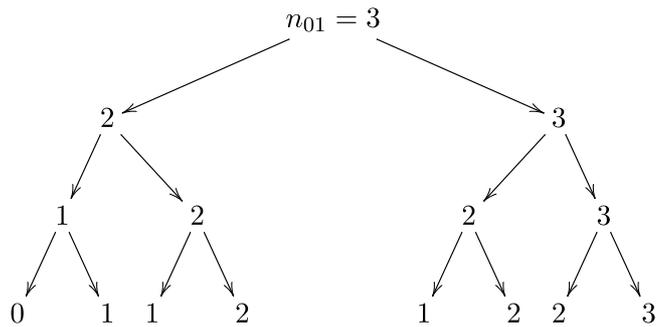


Схема 1

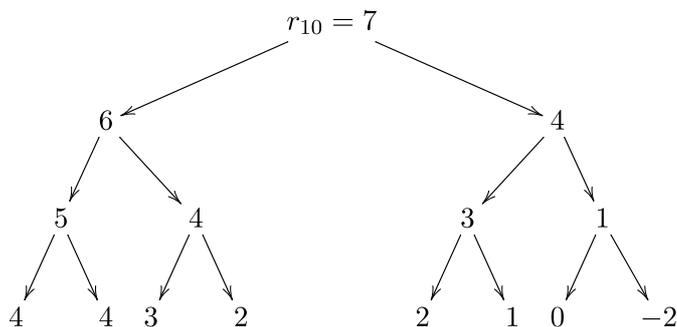


Схема 2

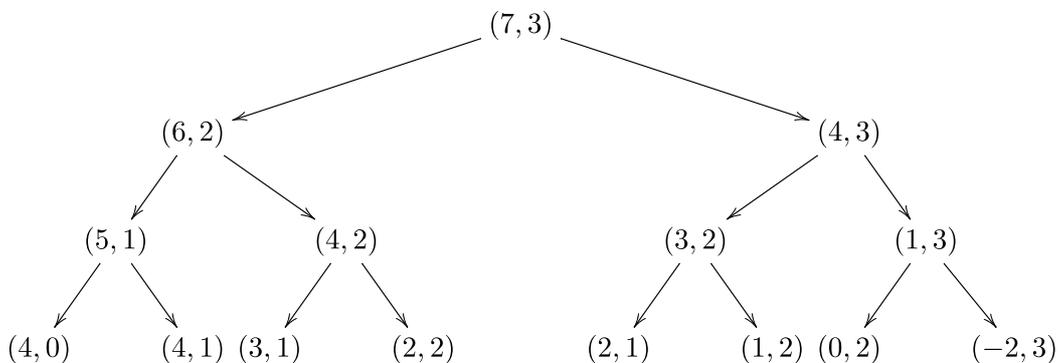


Схема 3

Рис. 4. Пример вычислений

Тогда приходим к ситуации, изображенной на рис. 4.

Теперь вычисляем искомое число N по (3) как сумму чисел, соответствующих траекториям (слева направо) (111), (110), (101), (100),

(011), (010), (001), (000), определенных свойствами 1а–4а:

$$N = 0 + 1 + 1 + 1 + 1 + 0 + 0 + 0 = 4,$$

что совпадает с результатом примера 2.

Вычисления для $\{r_{3j}\}$

j	$(T_3)_2$	$\{p_k - 1\}$	$l_k = 2^{p_k-1} - (T_{p_k-1})_{10}$	$r_{31} = r - S_{2^m-j} - \sum_{k=1}^{L_m} n_{p_k-1, l_k}$	r_{31}
1	111			$r - S_{111} = 7 - 3 = 4$	4
2	110	$p_1 - 1 = 2$	$2^2 - (T_{11})_{10} = 4 - 3 = 1$	$r - S_{110} - n_{21} = 7 - 2 - 1 = 4$	4
3	101	$p_1 - 1 = 1$	$2^1 - (T_1)_{10} = 2 - 1 = 1$	$r - S_{101} - n_{11} = 7 - 2 - 2 = 3$	3
4	100	$p_1 - 1 = 1$ $p_2 - 1 = 2$	$2^1 - (T_1)_{10} = 2 - 1 = 1$ $2^2 - (T_{10})_{10} = 4 - 2 = 2$	$r - S_{100} - n_{11} - n_{22} = 7 - 1 - 2 - 2 = 2$	2
5	011	$p_1 - 1 = 0$	$2^0 = 1$	$r - S_{011} - n_{01} = 7 - 2 - 3 = 2$	2
6	010	$p_1 - 1 = 0$ $p_2 - 1 = 2$	$2^0 = 1$ $2^2 - (T_{01})_{10} = 3$	$r - S_{010} - n_{01} - n_{23} = 7 - 3 - 2 = 2$	2
7	001	$p_1 - 1 = 0$ $p_2 - 1 = 1$	$2^0 = 1$ $2^1 - (T_0)_{10} = 2$	$r - S_{001} - n_{01} - n_{12} = 7 - 1 - 3 = 0$	0
8	000	$p_1 - 1 = 0$ $p_2 - 1 = 1$ $p_3 - 1 = 2$	$2^0 = 1$ $2^1 - (T_0)_{10} = 2$ $2^2 - (T_{00})_{10} = 4$	$r - S_{000} - n_{01} - n_{12} - n_{24} = 7 - 0 - 3 - 3 - 3 = -2$	-2

Теперь проведем вычисление искомого числа N в примере по формулам (3)–(9). Вначале вычислим все значения $\{n_{ij}\}$ для $i \in \{1, \dots, m\} = \{1, \dots, 3\}$, $j \in \{1, \dots, 2^i\}$ по формуле (7):

$$\begin{aligned} \{n_{01}\} &= \{3\}; \\ \{n_{1j}\} &= \{n_{11} = 3 - S_1 = 2; \\ &\quad n_{12} = 3 - S_0 = 3\} = \{2; 3\}; \\ \{n_{2j}\} &= \{n_{21} = 3 - S_{11} = 0; \\ &\quad n_{22} = 3 - S_{10} = 2; \\ &\quad n_{23} = 3 - S_{01} = 2; \\ &\quad n_{24} = 3 - S_{00} = 3\}; \\ \{n_{3j}\} &= \{n_{31} = 3 - S_{111} = 0; \\ &\quad n_{32} = 3 - S_{110} = 1; \\ &\quad n_{33} = 3 - S_{101} = 1; \\ &\quad n_{34} = 3 - S_{100} = 2; \\ &\quad n_{35} = 3 - S_{011} = 1; \\ &\quad n_{36} = 3 - S_{010} = 2; \\ &\quad n_{37} = 3 - S_{001} = 2; \\ &\quad n_{38} = 3 - S_{000} = 3\}, \end{aligned}$$

что совпадает со значениями $\{n_{ij}\}$ в схемах 1 и 3. Далее по формуле (8) вычислим значения $\{r_{3j}\}$, предварительно определив для каждого $j = 1, \dots, 8$ все входящие в формулу (8) значения, а именно $\{j\}$, $(T_3)_2$, $\{p_k\}$, $\{l_k\}$. Представим все вычисления для $\{r_{3j}\}$ в таблице с использованием вычисленных выше значений $\{n_{ij}\}$ для $i = 1, \dots, 3$. Итак, получаем, что $\{r_{31}\} = \{4, 4, 3, 2, 2, 1, 0, -2\}$, что совпадает с результатом непосредственного вычисления $\{r_{3j}\}$ по схемам 2 и 3 и с результатами применения рекуррентного соотношения (2) в примере 2. Далее по (3) и свойствам 1а–4а для $N(r, n) = N(7, 3)$ получаем, что

$$N = 0 + 1 + 1 + 1 + 1 + 0 + 0 + 0 = 4,$$

что совпадает с результатом прямого перебора всех вариантов размещения, приведенных в примере 2.

3. ВЕРОЯТНОСТНЫЙ АНАЛИЗ СХЕМЫ РАЗМЕЩЕНИЯ НЕРАЗЛИЧИМЫХ ЧАСТИЦ ПО НЕРАЗЛИЧИМЫМ ЯЧЕЙКАМ

3.1. Случайный процесс последовательного размещения частиц по ячейкам в схеме

Построим случайный процесс, исходы которого на i -м шаге описываются компонентами вектора $\bar{\mu}_i$, определенного во введении. Случайный процесс будет состоять в размещении на каждом шаге еще одной частицы по n неразличимым ячейкам. Состоянием процесса на i -м шаге будем считать так называемую вторую маркировку заполнений ячеек, то есть численности их заполнений всех возможных уровней в возрастающем порядке после размещения i частиц, представляемые вектором $\bar{\mu}_i$ при нулевом начальном заполнении.

Представим процесс в виде графа переходов из состояния в состояние пошагово по мере поединичного роста числа размещенных частиц, используя следующие обозначения: $E_j^{(i)}$ – j -е состояние после размещения i частиц, то есть на i -м шаге размещения. Состояния $\{E_j^{(i)}\}$ нумеруются в порядке попадания следующей частицы в ячейки с растущим предварительным уровнем заполнения (сначала – в одну из пустых ячеек, если они есть, затем – в одну из ячеек с одной частицей, если они есть, затем – с двумя частицами, если они есть, и так далее). Таким образом, состояния на каждом шаге будут упорядочены по принятому правилу. Вектор $\bar{\mu}_i$, описывающий состояние процесса на i -м шаге, будет состоять из $i + 1$ компоненты, задающих численности яче-

ек с уровнями заполнений от 0 до i в возрастающем порядке. Начальный вектор $\bar{\mu}_0 = (n)$ состоит из одной компоненты (все ячейки пусты). Для наглядности на ребрах графа укажем вероятности переходов из состояния в состояние (рис. 5).

3.2. Распределение вероятностей векторов $\{\bar{\mu}_r\}$ и вычисление чисел N и N^*

Зафиксируем $i = r$, то есть r -й шаг случайного процесса размещения r частиц по n ячейкам. Тогда в графе (см. рис. 5) получим на r -м шаге перечень всех упорядоченных по вышеприведенному правилу состояний $\{\bar{\mu}_r\}$, иными словами, всех размещений r неразличимых частиц по n неразличимым ячейкам. Их количество и есть число $N^* = N^*(r, n)$ исходов в общей схеме размещений. Число исходов $N = N(r, n)$ в схеме без пустых ячеек можно получить, выкидывая из N^* состояний общей схемы состояния с нулевой первой компонентой числа пустых ячеек. Таким образом, получаем метод вычисления чисел N и N^* путем перечисления состояний в графе случайного процесса последовательного размещения частиц до r -го шага.

Кроме этого, в результате упорядочения состояний в графе (см. рис. 5) решена задача перечислительной комбинаторики явного представления всех размещений r частиц по n ячейкам исходной схемы.

Распределение вероятностей векторов $\{\bar{\mu}_r\}$ вычисляется по графу (см. рис. 5) путем сложения по всем траекториям графа до данного состояния — вектора $\bar{\mu}_r$ — произведений вероятностей последовательных переходов, указанных на ребрах графа.

Пример 6. Пусть $r = 6$. Тогда

$$\begin{aligned} \mathbf{P}(E_1^{(4)}) &= \mathbf{P}(\bar{\mu}_4 = (n-4, 4, 0, 0, 0)) = \\ &= \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \frac{n-3}{n}, \\ \mathbf{P}(E_2^{(4)}) &= \mathbf{P}(\bar{\mu}_4 = (n-3, 2, 1, 0, 0)) = \\ &= \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \frac{3}{n} + \frac{n-1}{n} \cdot \frac{2}{n} \cdot \frac{n-2}{n}, \end{aligned}$$

и так далее, для всех пяти состояний.

Перебор траекторий до каждого конечного состояния в рассмотренном примере при $r = 6$ не вызывает затруднений, хотя для больших значений r решение этой проблемы требует отдельного рассмотрения. Обсудим общий поря-

док перебора траекторий до данного состояния при вычислении его вероятности по схеме на рис. 5. В силу единственности начального состояния для каждого конечного состояния предлагается выписывать траектории справа налево по обратным стрелкам в графе и перебирать их сверху вниз на каждом шаге.

4. ИССЛЕДОВАНИЕ СТОХАСТИЧЕСКОЙ МОДЕЛИ СХЕМЫ

4.1. Моделирование схемы размещения r неразличимых частиц по n неразличимым ячейкам

Способ 1. В п.3 теоретически найдено дискретное распределение вероятностей всех состояний схемы на любом шаге. Поэтому по полученному распределению на r -м шаге проводим моделирование состояний схемы методом маркировки (см. [1, гл. 3, §4]).

Замечание 5. Если при $r \geq n$ требуется смоделировать аналогичную схему без пустых ячеек, в графе перечисленных состояний на r -м шаге (см. рис. 5) исключаем состояния с нулевой компонентой μ_0 в векторе $\bar{\mu}_r$, а вероятности остальных состояний делим на единицу минус суммарная вероятность состояний с нулевой компонентой и далее применяем метод маркировки.

Способ 2. Известным, например, из [1, гл. 3, §1] способом моделируем заданное количество M_1 размещений r неразличимых частиц по n различимым ячейкам; выбираем из них M_2 повторяющихся, остальные маркируем по уровням заполнения в возрастающем порядке, выкидывая одинаковые. Получаем L_1 исходов вида вектора $\bar{\mu}_r$. С отбракованными M_2 размещениями повторяем ту же процедуру, что и с исходными M_1 размещениями, отбраковывая из них M_3 повторяющихся, и получаем L_2 исходов того же вида. Далее, заменяя M_i на M_{i+1} , а L_i на L_{i+1} , $i = 2, 3, \dots$, повторяем ту же процедуру k раз до тех пор, пока среди отбракованных вариантов не будет повторяющихся. Тогда, объединяя исходы вида вектора $\bar{\mu}_r$, получаем

$$L = \sum_{j=1}^k L_j$$

исходов требуемой схемы.

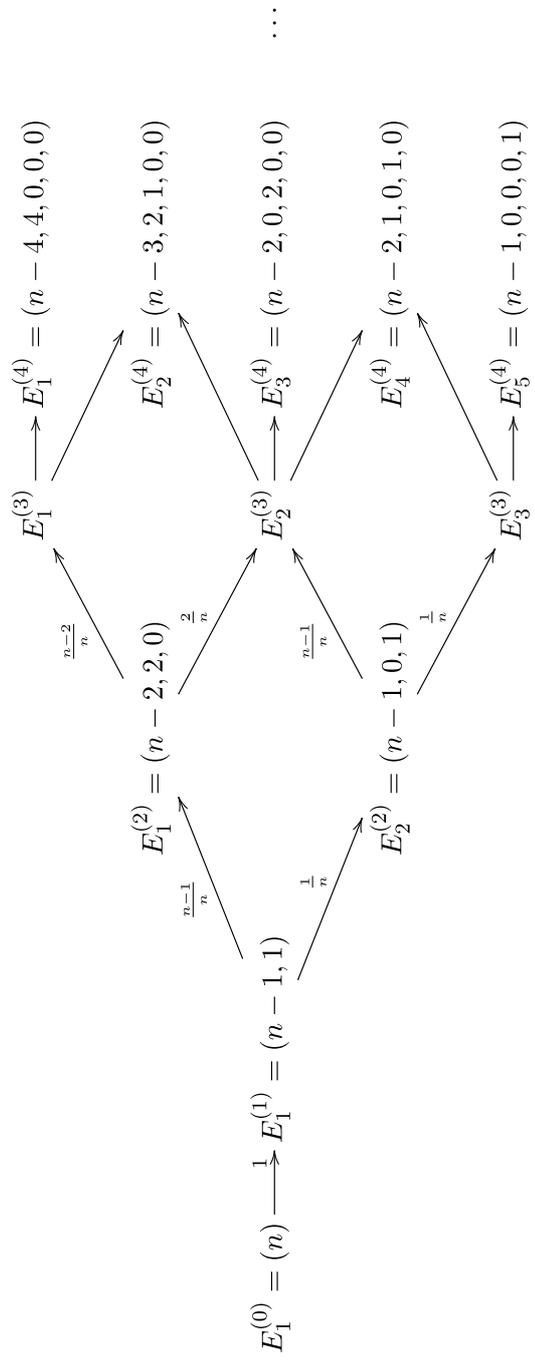


Рис. 5. Граф переходов

Замечание 6. Если при $r \geq n$ нужно смоделировать исходы аналогичной схемы без пустых ячеек, то известным, например, из [1, гл. 3, §1] способом моделируем M_1 размещений r неразличимых частиц по n различным ячейкам без пустых ячеек и применяем к ним второй способ.

4.2. Приближенное вычисление чисел N^* и N методом, использующим стохастическое моделирование

Воспользуемся вторым способом моделирования размещений соответственно в общей схеме, допускающей пустые ячейки с общим числом исходов N^* , и в схеме без пустых ячеек с общим числом исходов N . Тогда можно считать, что

$$N^* \binom{n+r-1}{r}^{-1} \approx \frac{L}{M_1},$$

откуда получаем, что

$$N^* \approx \frac{L}{M_1} \binom{n+r-1}{r}.$$

Аналогично, по замечанию 6, из того, что

$$N \binom{r-1}{n-1}^{-1} \approx \frac{L}{M_1},$$

находим приближенное значение N :

$$N \approx \frac{L}{M_1} \binom{r-1}{n-1}.$$

О МЕТОДЕ МАРКИРОВКИ

Для полноты изложения приведем краткие основные сведения о методе маркировки (см. [1]).

СВЕДЕНИЯ ОБ АВТОРАХ:

Энатская Наталья Юрьевна

доцент, к. ф.-м. н.
Московский институт электроники и математики
Национального исследовательского университета
«Высшая школа экономики»
ул. М. Пионерская, 12, 113054 Москва, Россия
эл. почта: nat1943@mail.ru

Хакимуллин Евгений Робертович

профессор, к. ф.-м. н.
Московский институт электроники и математики
Национального исследовательского университета
«Высшая школа экономики»
ул. М. Пионерская, 12, 113054 Москва, Россия
эл. почта: evgeni.hakimullin@mail.ru

Колчин Андрей Валентинович

к. ф.-м. н.
эл. почта: andrei.kolchin@gmail.com

Метод маркировки является одним из методов генерирования («разыгрывания») дискретной случайной величины с заданным законом распределения

$$\mathbf{P}(X = x_k) = p_k.$$

На отрезке $[0, 1]$ изобразим точки вида $\sum_{k=1}^s$, $s = 1, 2, \dots$

Пусть r — возможное значение случайной величины R , равномерно распределенной на отрезке $[0, 1]$, тогда

$$\begin{aligned} \mathbf{P}(0 < r < p_1) &= p_1, \\ \mathbf{P}(p_1 < r < p_1 + p_2) &= p_2, \dots, \\ \mathbf{P}\left(\sum_{i=1}^{k-1} p_i < r < \sum_{i=1}^k p_i\right) &= p_k, \dots \end{aligned}$$

Отсюда следует, что попадание случайного числа R на k -й отрезок моделирует полученное значение случайной величины $X = x_k$.

Замечание. Для многих основных распределений так называемый коэффициент воспроизводимости $\gamma_k = p_{k+1}/p_k$ имеет для всех k удобное общее выражение как функции от k . Поэтому в таком случае при использовании метода маркировки нет необходимости загрузки в память всего ряда распределения, вместо этого $\{p_k\}$ вычисляется по мере необходимости по формуле

$$p_{k+1} = \gamma_k p_k.$$

ЛИТЕРАТУРА

1. Энатская Н. Ю., Хакимуллин Е. Р. Стохастическое моделирование. МИЭМ, Москва, 2012. 185 с.

Enatskaya, Natalia

Moscow Institute of Electronics and Mathematics,
Higher School of Economics
12 M. Pionerskaya St., 113054 Moscow, Russia
e-mail: nat1943@mail.ru

Khakimullin, Evgeny

Moscow Institute of Electronics and Mathematics,
Higher School of Economics
12 M. Pionerskaya St., 113054 Moscow, Russia
e-mail: evgeni.hakimullin@mail.ru

Kolchin, Andrey

e-mail: andrei.kolchin@gmail.com