

БИНОМИАЛЬНЫЕ СИСТЕМЫ СЧИСЛЕНИЯ И НУМЕРАЦИИ КОМБИНАТОРНЫХ ОБЪЕКТОВ*)

А. Я. Зантен, ван

Изучается роль различных систем счисления в решении задач пересчисления комбинаторных объектов в лексикографическом порядке и порядке их минимального изменения, аналогичном известному коду Грея. Изложение иллюстрируется примерами.

Введение

Пусть V — множество некоторых комбинаторных объектов и $|V| = N$. Далее, пусть объекты V линейно упорядочены согласно некоторому критерию, определяющему относительный порядок объектов. Тогда каждому объекту можно приписать индекс или номер, который обычно изменяется от 1 до N или от 0 до $N - 1$.

Определение номера объекта обычно называется задачей нумерации. Обратная задача нахождения объекта по данному номеру называется задачей денумерации. Очень часто порядок на множестве V определяется рекурсивно. В таком случае требуется получить нерекурсивное решение задачи (де-)нумерации.

Поскольку многие комбинаторные объекты представляются в виде двоичных наборов или, в более общем случае, в виде слов в алфавите из q букв, $q \geq 2$, они могут быть упорядочены лексикографически, и тогда можно говорить о лексикографическом коде. Другим порядком является порядок минимального изменения, или код Грея, при котором изменение при переходе к следующему объекту является «минимальным».

Оказывается, что в большинстве случаев существуют некоторые системы счисления, которые очень подходят для решения задач нумерации [4]. Так, обычная двоичная система счисления является подходящей системой для лексикографически упорядоченного кода $L(n)$ двоичных наборов длины n . Это же верно и для кода Грея $G(n)$ тех же объектов.

*) Перевод с английского доклада автора, прочитанного на Международной Сибирской конференции по исследованию операций, которая состоялась в Новосибирске с 22 по 27 июня 1998 г.

Менее известными примерами являются нумерации двоичных наборов длины n одинакового веса k . Подходящей системой для лексикографического кода $L(n, k)$ является биномиальная система счисления, тогда как в случае кода Грея $G(n, k)$ удобно использовать так называемую чередующуюся биномиальную систему счисления [5]. Еще две биномиальные системы (называемые смешанными биномиальными системами) оказываются удобными для нумерации кода $C(n, k)$ разбиений числа n на k частей. В $C(n, k)$ эти разбиения представляются в виде $r_k r_{k-1} \dots r_1$, где $r_1 + r_2 + \dots + r_k = n$, $r_i \geq 0$, и упорядочены в некотором порядке минимального изменения [6].

Обобщением обычной биномиальной системы является циклономиальная система, которая может быть использована для нумерации лексикографического кода $L^q(n, k)$ таких наборов (a_1, \dots, a_n) в алфавите $\{0, 1, \dots, q-1\}$, что $\sum_{i=1}^n a_i = k$ (см. [2]).

1. Комбинации

Любая k -комбинация n различных предметов может быть представлена двоичным набором длины n с k единицами. Такие наборы могут быть упорядочены как лексикографически, так и согласно критерию минимального изменения, при котором каждый набор отличается от своего предшественника точно в двух разрядах.

Лексикографический код $L(n, k)$ определяется рекурсивно следующим образом:

$$\begin{aligned} L(n, 0) &= (00 \dots 0), & L(n, n) &= (11 \dots 1), \\ L(n, k) &= \begin{bmatrix} 0L(n-1, k) \\ 1L(n-1, k-1) \end{bmatrix}, & 0 < k < n. \end{aligned}$$

Кодовые слова нумеруются числами от 0 до $\binom{n}{k} - 1$. Подходящей системой счисления для такой нумерации является хорошо известная *биномиальная система счисления* (см., например, [1]), основанная на следующем свойстве.

Утверждение 1. Пусть $k \geq 1$ — произвольное целое. Любое целое число $N \geq 0$ единственным образом представимо в виде

$$N = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \binom{a_{k-2}}{k-2} + \dots + \binom{a_1}{1},$$

где $a_k > a_{k-1} > a_{k-2} > \dots > a_1 \geq 0$.

В этом случае используем запись: $N = (a_k a_{k-1} a_{k-2} \dots a_1)_s$. Номер i -го кодового слова l_i кода $L(n, k)$ можно представить в виде

$$i = (b_k b_{k-1} \dots b_1)_s,$$

где b_1, b_2, \dots, b_k — номера позиций единиц в l_i (справа налево), $0 \leq b_j \leq n-1$ для всех j .

Эти же слова могут быть упорядочены как код Грея $G(n, k)$ следующим образом:

$$G(n, 0) = (00 \dots 0), \quad G(n, n) = (11 \dots 1),$$

$$G(n, k) = \begin{bmatrix} 0G(n-1, k) \\ 1G(n-1, k)^R \end{bmatrix}, \quad 0 < k < n.$$

Обозначение $G(n-1, k-1)^R$ используется для списка $G(n-1, k-1)$, заданного в обратном порядке. Соответствующей системой счисления является *чередующаяся система счисления*, введенная в [5].

Утверждение 2. Пусть $k \geq 1$ — произвольное натуральное число. Любое $N \geq 0$ при четном k и любое $N > 0$ при нечетном k однозначно представимо в виде

$$N = \binom{a_k}{k} - \binom{a_{k-1}}{k-1} + \binom{a_{k-2}}{k-2} - \dots \pm \binom{a_1}{1},$$

где $a_k > a_{k-1} > a_{k-2} > \dots > a_1 \geq 0$.

В этом случае будем писать $N = (a_k a_{k-1} a_{k-2} \dots a_1)_a$. Номер i -го кодового слова g_i кода $G(n, k)$ представим в виде

$$i = (b_k + 1, b_{k-1} + 1, \dots, b_1 + 1)_a + \zeta_k,$$

где $\zeta_k = 0$ при четном k и $\zeta_k = -1$ при нечетном k .

2. Разбиения

Пусть $n \geq 0$ и $k > 0$. Под *разбиением числа n на k частей* понимается упорядоченное представление

$$n = r_1 + r_2 + \dots + r_k,$$

где r_1, r_2, \dots, r_k — целые неотрицательные числа. Это разбиение будем записывать в виде слова

$$r_k r_{k-1} \dots r_1. \quad (1)$$

Разбиение (1) можно интерпретировать в терминах размещения частиц по ячейкам. Пусть имеется k помеченных ячеек и n неразличимых предметов. Тогда каждому разбиению (1) соответствует единственное размещение n частиц по ячейкам. Список, возникающий при упорядочивании слов (1) в лексикографическом порядке, обозначим через $D(n, k)$. Слова из $D(n, k)$ занумерованы числами от 0 до $\binom{n+k-1}{n} - 1$. Если d_i является i -м словом в $D(n, k)$, то

$$i = \binom{n+k-1}{n} - 1 - (a_{k-1}, a_{k-2}, \dots, a_1)_s,$$

где

$$a_l = \sum_{j=1}^l r_j + l - 1, \quad 1 \leq l \leq k - 1.$$

Рассмотрим упорядоченность разбиений (1) согласно коду Грея. Под этим мы понимаем, что список всех разбиений упорядочен так, что каждое размещение частиц по ячейкам получается из предшествующего перемещением единственной частицы из какой-либо ячейки в другую. Определим код Грея $C(n, k)$ следующим образом (что является небольшой вариацией из [5]):

$$C(n, 1) = (n),$$

$$C(n, k) = \begin{bmatrix} 0C(n, k-1) \\ 1C(n-1, k-1)^R \\ 2C(n-1, k-1) \\ 3C(n-3, k-1)^R \\ \vdots \end{bmatrix}, \quad k > 1.$$

Для решения задачи нумерации $C(n, k)$ понадобится следующее

Утверждение 3. Пусть $k \geq 1$ — произвольное натуральное число. Произвольное $N \geq 0$ однозначно представимо в виде

$$N = \varepsilon_k \binom{a_k}{k} + \varepsilon_{k-1} \binom{a_{k-1}}{k-1} + \varepsilon_{k-2} \binom{a_{k-2}}{k-2} + \cdots + \varepsilon_1 \binom{a_1}{1} + \varepsilon_0 \binom{a_0}{0} + \zeta,$$

где $a_k > a_{k-1} > a_{k-2} > \cdots > a_1 > a_0$, $a_0 \in \{-1, 0\}$, $\varepsilon_k = 1$ и $\varepsilon_i \in \{-1, 1\}$ при $0 \leq i \leq k-1$, причем $\varepsilon_i = \varepsilon_{i+1}$ тогда и только тогда, когда $\text{par}(a_i) \neq \text{par}(a_{i+1})$. Число ζ равно -1 при $\text{par}(k) = \text{par}(a_k)$ и равно 0 при $\text{par}(k) \neq \text{par}(a_k)$.

Четность любого числа a определяем следующим образом: $\text{par}(a) = 1$, если a четно, и $\text{par}(a) = -1$, если a нечетно. Кроме того, полагаем, как обычно, $\binom{0}{0} = 1$ и $\binom{-1}{0} = 0$, что необычно.

Для любого $k \geq 1$ утверждение 3 позволяет представлять единственным образом любое число $N \geq 0$ при нечетном k и любое число $N > 0$ при четном k в виде $N = (a_k a_{k-1} \dots a_0)_e$, где a_k четно, $a_k > a_{k-1} > \cdots > a_0$ и $a_0 \in \{-1, 0\}$.

Аналогичное представление $N = (a_k a_{k-1} \dots a_0)_o$ в случае нечетного a_k имеем для любого $N \geq 0$ при четном k и для любого $N > 0$ при нечетном k .

В обоих случаях для N имеем

$$N = \varepsilon_k \binom{a_k}{k} + \varepsilon_{k-1} \binom{a_{k-1}}{k-1} + \cdots + \varepsilon_0 \binom{a_0}{0},$$

где $\varepsilon_k = 1$ и $\varepsilon_i = \varepsilon_{i+1}$, если $\text{par}(a_i) \neq \text{par}(a_{i+1})$, и $\varepsilon_i = -\varepsilon_{i+1}$, если $\text{par}(a_i) = \text{par}(a_{i+1})$. Эти системы счисления называются «четная смешанная биномиальная система счисления» и «нечетная смешанная биномиальная система счисления».

В дальнейшем нам понадобятся частичные суммы

$$a_0 = 0, \quad a_j = \sum_{i=1}^r r_i, \quad 1 \leq j \leq k.$$

Введем обозначения:

$$\delta_j = \begin{cases} 1, & \text{если } r_j \text{ нечетно,} \\ 0, & \text{если } r_j \text{ четно,} \end{cases} \quad 1 \leq j \leq k$$

и

$$b_j := a_j + j - 1 + \delta_{j+1}, \quad 0 \leq j \leq k - 1.$$

Пусть c_i является i -м кодовым словом кода $C(n, k)$. Если $n + k$ четно, то номер i может быть представлен в четной смешанной системе счисления

$$i = \binom{n+k-1}{n} - 1 - (b_{k-1}b_{k-2}\dots b_0)_e - \zeta,$$

где $\zeta = 0$ при четном n и $\zeta = -1$ при нечетном n .

Подобное выражение в нечетной смешанной системе счисления (e вместо o) верно, если $n + k$ нечетно.

3. Примеры

(i) Если слово 110101 рассмотреть как слово из $L(6, 4)$, то его номер представляется в виде $(5420)_s$, что соответствует величине $\binom{5}{4} + \binom{4}{3} + \binom{2}{2} + \binom{0}{1} = 5 + 4 + 1 + 0 = 10$.

Если слово 110101 рассмотреть как слово из $G(6, 4)$, то его номером окажется число $(6531)_a = \binom{6}{4} - \binom{5}{3} + \binom{3}{2} - \binom{1}{1} = 15 - 10 + 3 - 1 = 7$.

(ii) Обратно, пусть требуется вычислить $l_6 \in L(6, 4)$. Записывая

$$6 = \binom{5}{4} + \binom{3}{3} + \binom{1}{2} + \binom{0}{1} = (5310)_s,$$

получаем $l_6 = 101011$.

Чтобы получить $g_6 \in G(6, 4)$, вычисляем $6 - \zeta_4 = 6$ в чередующейся биномиальной системе счисления для $k = 4$, получая

$$6 = \binom{6}{4} - \binom{5}{3} + \binom{3}{2} - \binom{2}{1} = (6532)_a.$$

Поэтому кодовый набор имеет единицы в 5, 4, 2 и 1-й позициях. Следовательно, $g_5 = 110110$.

(iii) Рассмотрим разбиение $c_i := 1201 \in C(5, 4)$.

Последовательно вычисляем

$$\bar{a} = (a_4, a_3, a_2, a_1, a_0) = (5, 4, 2, 2, 0),$$

$$\bar{\varepsilon} = (\varepsilon_3, \varepsilon_2, \varepsilon_1, \varepsilon_0) = (1, -1, -1, -1),$$

$$\bar{\delta} = (\delta_4, \delta_3, \delta_2, \delta_1) = (1, 0, 0, 0, 0).$$

С помощью этих параметров находим $b_3 = 4 + 2 + 1 = 7$, $b_2 = 2 + 1 + 0 = 3$, $b_1 = 2 + 0 + 0 = 2$, $b_0 = 0 + (-1) + 0 = -1$. Так как числа $n + k = 9$ и $n = 5$ нечетны, то номер разбиения 1 2 0 2 равен

$$\binom{8}{5} - 1 - (732(-1))_o + 1 = \binom{8}{5} - \left[\binom{7}{3} - \binom{3}{2} - \binom{2}{1} - \binom{-1}{0} \right] = 26.$$

(iv) Для обратной задачи денумерации используем следующее правило, которое может быть легко выведено из результатов предыдущего раздела.

Пусть i — номер некоторого слова $r_k r_{k-1} \dots r_1 \in C(n, k)$ и пусть $(b_{k-1} b_{k-2} \dots b_0)_p$ ($p = e$ или $p = o$) является представлением числа $\binom{n+k-1}{n} - 1 - \zeta - i$. Тогда при любом i , $1 \leq i < k$, имеем $r_i = b_i - b_{i-1} - \delta_{i+1} + \delta_i - 1$, где $\delta_i \in \{0, 1\}$, причем $\delta_1 = 1$ при $par(b_0) = 1$ и $\delta_i = 1$ при $1 < i < k$ и $par(b_{i-1}) = par(b_{i-2})$.

Применим это правило для нахождения $c_{16} \in C(5, 4)$. Так как числа $n + k = 9$ и $n = 5$ нечетны, то представляя число $55 - 16 + 1 = 40$ в нечетной биномиальной системе счисления, находим

$$40 = (7420)_o.$$

Следовательно, $\bar{b} = (b_3, b_2, b_1, b_0) = (7, 4, 2, 0)$ и $\bar{\delta} = (\delta_4, \delta_3, \delta_2, \delta_1) = (0, 1, 1, 1)$. Поэтому $r_3 = 7 - 4 - 0 + 1 = 3$, $r_2 = 4 - 2 - 1 + 1 - 1 = 1$ и $r_1 = 2 - 0 - 1 + 1 - 1 = 1$. Таким образом, искомым кодовым словом является 0311.

ЛИТЕРАТУРА

1. Кнут Д. Искусство программирования для ЭВМ: В 2 т. М.: Мир, 1976. Т. 1: Основные алгоритмы.
2. Roos C., Snijders A., Zanten A. J. van. Cyclonomial number systems and the ranking of lexicographically ordered constant-sum codes // Arc Combinatoria. 1998. V. 48. P. 173-184.
3. Wilf H. S. Combinatorial algorithms: an update. Philadelphia, PA: SIAM, 1989. VI+47 p. (CBMS-NSF Regional Conf. Ser. in Appl. Math.; V. 55).

4. **Zanten A. J. van.** Minimal-change codes and number systems. Delft: Univ. of Technology, Fac. of Mathematics and Informatics, 1991. Rept. 91-29.
5. **Zanten A. J. van.** Index system and separability of constant weight Gray codes // IEEE Trans. on Inform. Theory, 1991. V. 37, N 4. P. 1229–1233.
6. **Zanten A. J. van.** The ranking problem of a Gray code for compositions // Ars Combinatoria. 1995. V. 41. P. 257–268.

Адрес автора:

A. J. van Zanten

Delft University of Technology,
Department of Mathematics,
P. O. Box 5031, 2600 GA Delft,
The Netherlands.

E-mail:

a.j.vanzanten@twi.tudelft.nl

Статья поступила

3 сентября 1998 г.