

БЫСТРОДЕЙСТВУЮЩИЙ МЕТОД БИНОМИАЛЬНОГО НУМЕРАЦИОННОГО КОДИРОВАНИЯ

Предлагается быстродействующий метод биномиального нумерационного кодирования на основе комбинаторного алгоритма вычисления номера двоичной комбинации, использующего рекуррентные последовательности биномиальных коэффициентов. Приведено теоретическое обоснование правильности разработанного алгоритма биномиальной нумерации, а также его блок-схема. Рассматривается пример вычисления номера двоичной комбинации согласно быстродействующему методу биномиального нумерационного кодирования.

1. Введение

Методы нумерационного кодирования позволяют сжимать данные, вероятностные характеристики которых являются неизвестными. Это положительное качество выгодно отличает нумерационное кодирование от методов статистического сжатия, которым для эффективной работы необходимо, как правило, проведение предварительной оценки информации и составление таблиц условных вероятностей. Кроме того, получаемый при нумерации равномерный код, во-первых, упрощает меры синхронизации работы передающего и приемного устройств, а, во-вторых, является более устойчивым к сериям ошибочных изменений значений разрядов. Приведенные достоинства обеспечивают достаточно широкое распространение методов нумерационного кодирования, определяют актуальность и перспективу дальнейшей их разработки [1, 2].

Эффективность метода нумерации обеспечивается:

1) “разумными” ограничениями на элементы нумеруемых последовательностей, поскольку от содержания таких ограничений напрямую зависят существование и распространность нумеруемых последовательностей;

2) объемом вычислительных затрат при нахождении номера последовательности.

Особое место среди методов нумерации двоичных последовательностей занимает метод биномиального нумерационного кодирования, основанный на применении числовой функции биномиальной системы счисления [3, 4]. Практическую значимость данного метода обуславливает тот факт, что нумеруемыми последовательностями являются равновесные комбинации, поскольку в качестве ограничения рассматривается равенство количества двоичных единиц заданному фиксированному значению. Равновесный код широко используется в системах передачи информации, управления и шифрования данных. Более того, любую двоичную последовательность можно представить как равновесную при известном значении числа содержащихся в ней единиц. Таким образом, биномиальное нумерационное кодирование позволяет перечислять (сжимать) любые двоичные последовательности и генерировать любой равновесный код, содержащий заданное число единиц.

Следует отметить, что приведенными достоинствами биномиальной нумерации (сжатия) эффективно воспользоваться можно только в том случае, если скорость проведения вычислений номера будет достаточно высокой, а объем вычислительных затрат при этом небольшим и приемлемым с точки зрения требований аппаратно-программной реализации. Основной вклад в сложность вычислений номера вносит поиск значений весовых коэффициентов биномиальной числовой функции. В качестве весовых коэффициентов используются числа сочетаний, в параметрах которых присутствуют значения n – длины нумеруемой последовательности и k – количества содержащихся в ней двоичных единиц. Нахождение точных значений чисел сочетаний, как это требует числовая функция биномиальной системы счисления, при достаточно больших n и k сталкивается с определенными трудностями, основными из которых являются значительное время вычислений и переполнение разрядной сетки в связи с огромными значениями факториалов чисел. При этом следует отметить, что наиболее полно преимущества биномиального нумерационного кодирования

раскрываются при нумерации длинных последовательностей. Таким образом, для дальнейшего развития метода биномиального нумерационного кодирования важными являются задачи:

- увеличение скорости преобразования двоичной последовательности в номер;
- упрощение вычисления номера и, как следствие, снижение объема аппаратно-программных затрат при практической реализации.

В качестве одного из решений указанных выше задач предлагается использование рекуррентного алгоритма вычисления весовых коэффициентов биномиальной числовой функции. Предлагаемый алгоритм не требует применения сложных цифровых устройств при технической реализации и является весьма эффективным с точки зрения скорости его работы.

2. Метод биномиального нумерационного кодирования

В соответствии с методом биномиального нумерационного кодирования нумеруемыми (сжимаемыми) последовательностями являются последовательности $Y = (y_{n-1} \dots y_i \dots y_1 y_0)$,

сумма двоичных разрядов y_i которых удовлетворяет ограничению $\sum_{i=0}^{n-1} y_i = k$.

Такие двоичные последовательности представляют равновесный код с параметрами n – длина кодовой комбинации и k – количество содержащихся в ней единиц. При этом следует отметить, что любая двоичная последовательность может быть рассмотрена как равновесная при известном значении количества k составляющих ее единиц.

Переход от исходной последовательности $Y = (y_{n-1} \dots y_i \dots y_1 y_0)$ к соответствующему номеру F осуществляется в два этапа, на первом из которых производится преобразование $Y = (y_{n-1} \dots y_i \dots y_1 y_0)$ к биномиальному числу $X = (x_{r-1} \dots x_i \dots x_1 x_0)$, а на втором – преобразование биномиального числа $X = (x_{r-1} \dots x_i \dots x_1 x_0)$ в номер F . Первое преобразование выполняется в соответствии с системами ограничений для биномиальных чисел X [3, 4]:

$$\begin{cases} k \leq r \leq n-1 \\ q = k \\ x_0 = 1 \end{cases} \quad \text{и} \quad \begin{cases} n-k = r-q \\ 0 \leq q \leq k-1 \\ x_0 = 0 \end{cases}, \quad (1)$$

где r – длина биномиального числа X , а q – количество содержащихся в X двоичных единиц. Анализ ограничений (1) показывает, что первое преобразование состоит в отбрасывании в двоичной последовательности Y единиц справа до первого появления нуля или отбрасывании нулей справа до первого появления единицы.

Второе преобразование, заключающееся в получении номера F из биномиального числа $X = (x_{r-1} \dots x_i \dots x_1 x_0)$, осуществляется на основании числовой функции двоичной биномиальной системы счисления [3, 4]:

$$F = x_{r-1} C_{n-1}^{k-q_{r-1}} + \dots + x_i C_{n-r+i}^{k-q_i} + \dots + x_0 C_{n-r}^{k-q_0} = \sum_{i=0}^{r-1} x_i C_{n-r+i}^{k-q_i}, \quad (2)$$

где q_i – сумма единичных значений разрядов x_i от $(r-1)$ -го разряда до $(i+1)$ -го включительно:

$$q_i = \sum_{t=i+1}^r x_t, \quad (3)$$

здесь $q_{r-1} = x_r = 0$.

Весовой коэффициент i -го разряда в числовой функции (2) представляет собой число сочетаний $C_{n-r+i}^{k-q_i}$, которое зависит как от позиции $i = 0, 1, \dots, r-1$ рассматриваемого разряда, так и от суммы q_i предшествующих этому разряду двоичных значений x_i .

Общий алгоритм биномиального нумерационного кодирования двоичной последовательности длины n , содержащей количество k единиц, представлен на рис. 1.



Рис. 1. Общий алгоритм биномиального нумерационного кодирования

Разработаны различные способы аппаратно-программной реализации биномиального нумерационного кодирования двоичных последовательностей, отличающиеся друг от друга такими характеристиками, как быстродействие и объем аппаратно-программных затрат [3, 4].

Наиболее высоким быстродействием обладает табличный способ нахождения номеров, где в качестве адреса (индекса) в таблице служит двоичная последовательность, которой соответствует ячейка с требуемым номером. Номера заранее заносятся в таблицу согласно биномиальной числовой функции. Данный способ реализуется с помощью устройств памяти. На вход устройства памяти подаются равновесные комбинации, а на выходе устанавливаются соответствующие номера. Однако с ростом длины двоичных последовательностей объем памяти, а следовательно, и аппаратно-программные затраты при реализации табличного способа резко возрастают.

Другим способом реализации рассматриваемого метода кодирования является способ последовательного перебора двоичных равновесных последовательностей, который обладает наименьшим быстродействием. Согласно этому способу последовательно перебираются равновесные комбинации, организованные в лексикографическом порядке, и одновременно с ними перебираются двоичные номера. Практическая реализация данного способа основывается на двух счетчиках, один из которых является биномиальным, а другой – обычным двоичным счетчиком. Синтез структуры биномиального счетчика осуществляется в соответствии с биномиальной числовой функцией. При относительно больших значениях длины равновесного кода количество всех возможных кодовых комбинаций становится весьма значительным, в связи с чем резко снижается скорость биномиальной нумерации кодовых последовательностей.

С точки зрения высокого быстродействия и малого объема аппаратно-программных затрат оптимальным способом реализации метода биномиального нумерационного кодирования можно считать способ, использующий комбинаторные рекуррентные соотношения для весовых коэффициентов числовой функции (2), в качестве которых служат числа сочетаний, поскольку именно нахождение чисел сочетаний требует значительных времени и объема аппаратно-программных затрат.

3. Вычисление биномиальных коэффициентов

Основными причинами, усложняющими процесс нахождения биномиальных коэффициентов C_n^k при больших значениях параметров n и k , являются:

- 1) огромные значения факториалов $n!$ и $k!$, используемых при вычислении C_n^k ;
- 2) большое количество выполняемых операций умножения и деления.

Избежать огромных значений факториалов, исключить операции деления и существенно снизить количество операций умножения позволяет метод вычисления биномиальных коэффициентов на основе канонического разложения чисел [5]. Метод основывается на разложении чисел сочетаний, а также факториалов их параметров на простые числа, возведенные в определенные степени. При этом операции умножения и деления сводятся к арифметическим операциям над степенями простых чисел. Недостатком данного метода является необходимость хранить или вычислять все простые числа до значения нижнего параметра числа сочетаний.

При программной реализации вычисления C_n^k часто используют алгоритм на основе рекуррентного соотношения

$$C_n^k = C_n^{n-k} = \frac{n-k+1}{k} \cdot C_n^{k-1}. \quad (4)$$

Данный метод является достаточно простым, однако при его аппаратной реализации возникают трудности, связанные с необходимостью выполнять операции умножения и деления над большими числами.

В работе [6] предлагается метод вычисления биномиальных коэффициентов при помощи последовательностей чисел сочетаний, в котором не применяются операции умножения и деления. В основу данного метода положены формулы рекуррентного разложения биномиальных коэффициентов следующего вида:

$$C_n^k = C_{n-1}^k + C_{n-2}^{k-1} + \dots + C_{n-k-1}^0, \quad (5)$$

$$C_n^k = C_{n-1}^{k-1} + C_{n-2}^{k-1} + \dots + C_{k-1}^{k-1}. \quad (6)$$

При вычислении C_{i+j}^j создается i -я последовательность биномиальных коэффициентов, содержащая $j+1$ элементов, где $i = 0, 1, \dots, n-k$, а $j = 0, 1, \dots, k$. Число сочетаний C_{i+j}^j при этом представляется как j -й элемент, расположенный в i -й последовательности. В случае необходимости определить коэффициент $C_{(i+1)+1}^j$ строится новая $(i+1)$ -я последовательность на основе i -й предшествующей. В общем виде i -я последовательность будет выглядеть как

$$C_i^0 \dots C_{i+j}^j \dots C_{i+k}^k. \quad (7)$$

При $i = n-k$ и $j = k$ число сочетаний $C_{i+j}^j = C_{n-k+k}^k = C_n^k$, т.е. k -й элемент $(n-k)$ -й последовательности представляет собой искомый биномиальный коэффициент C_n^k .

Начальная последовательность $i=0$ состоит из $(k+1)$ единиц согласно $C_0^0 = C_1^1 = \dots = C_k^k = 1$. Элементы следующей последовательности $i=1$ в соответствии с формулой (5) получаются в процессе суммирования единиц нулевой последовательности, начиная с самой левой и до номера искомого элемента включительно, который представляет собой значение верхнего параметра искомого числа сочетаний. Аналогично может быть получена вторая, третья и т.д. последовательности.

Для упрощения вычисления элементов новой последовательности можно воспользоваться значениями уже вычисленных элементов формируемой последовательности.

Теорема 1. В последовательностях биномиальных коэффициентов, построенных согласно рекуррентному разложению (5), сумма значений j -го элемента $(i-1)$ -й последовательности и $(j-1)$ -го элемента i -й последовательности представляет собой значение j -го элемента i -й последовательности:

$$C_{i+j}^j = C_{(i-1)+j}^j + C_{i+(j-1)}^{j-1}. \quad (8)$$

Доказательство. В выражении (8) перегруппируем переменные i и j в нижних параметрах чисел сочетаний: $C_{i+j}^j = C_{(i+j)-1}^j + C_{(i+j)-1}^{j-1}$. Введя переменную z вместо суммы $(i+j)$, приходим к известному правилу Паскаля: $C_z^j = C_{z-1}^j + C_{z-1}^{j-1}$. Теорема 1 доказана.

В результате для вычисления j -го элемента i -й последовательности биномиальных коэффициентов необходимо просуммировать j -й элемент предшествующей $(i-1)$ -й последовательности с $(j-1)$ -м элементом искомой i -й последовательности.

Пример. Вычислить значение биномиального коэффициента C_{10}^4 с помощью последовательности чисел сочетаний.

Решение. Поскольку номер искомого элемента в последовательности чисел сочетаний $j=4$, а сумма номеров строки и элемента $i+j=10$, то в соответствии с методом вычисления номер последовательности $i=6$. Таким образом, $j=\overline{0,4}$, $i=\overline{0,6}$ и, следовательно, необходимо построить семь последовательностей, состоящих из пяти биномиальных коэффициентов.

Последовательности биномиальных коэффициентов, получившиеся в результате применения метода для заданных параметров $n=10$ и $k=4$, сведены в табл. 1.

Таблица 1. Пример вычисления C_{10}^4

$i \backslash j$	0	1	2	3	4
0	1	1	1	1	1
1	1	2	3	4	5
2	1	3	6	10	15
3	1	4	10	20	35
4	1	5	15	35	70
5	1	6	21	56	126
6	1	7	28	84	210

Число, расположенное на пересечении четвертого столбца и шестой строки, соответствует значению биномиального коэффициента с параметрами $n=10$ и $k=4$, т.е. $C_{10}^4 = 210$.

Рассмотренный метод вычисления биномиальных коэффициентов имеет следующие достоинства:

1) достаточно простая аппаратная реализация метода, поскольку содержит только операции сложения;

2) удобство применения для нумерационной функции (2), поскольку при вычислении биномиального коэффициента производится расчет биномиальных коэффициентов с меньшими значениями параметров, что соответствует порядку определения слагаемых суммы (2). Данная особенность позволяет ускорить процесс определения номера исходной последовательности.

4. Комбинаторный алгоритм вычисления номера двоичной комбинации

На основе общего алгоритма биномиального нумерационного кодирования, представленного на рис. 1, строится комбинаторный алгоритм вычисления номера двоичной комбинации.

Как следует из системы ограничений (1), преобразование равновесной комбинации Y в r -разрядное биномиальное число X (преобразование 1) может выполняться двумя способами: 1) отбрасыванием подряд расположенных $(n-r)$ нулей, начиная с $y_0=0$, до появления первой единицы, которая становится самым младшим единичным разрядом $x_0=1$ числа X ; 2) отбрасыванием подряд расположенных $(n-r)$ единиц, начиная с $y_0=1$, до появления первого нуля, который становится самым младшим нулевым разрядом $x_0=0$ числа X .

Переход от биномиального числа X к номеру F равновесной комбинации (преобразование 2) осуществляется на основании числовой функции биномиальной системы счисления (2). Значения биномиальных коэффициентов, входящих в состав числовой функции (2), вычисляются при помощи последовательности чисел сочетаний (7).

Теорема 2. Номер разряда i двоичного биномиального числа $X = (x_{r-1} \dots x_1 x_0)$, $i = 0, r-1$, связан с номером разряда $s \geq n-r$ двоичной равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$, имеющей число k единиц, следующим соотношением

$$i = s - n + r. \quad (9)$$

Доказательство. Двоичная равновесная комбинация $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$ характеризуется количеством n составляющих ее разрядов и числом k содержащихся в ней единиц. Получение r -разрядного биномиального числа $X = (x_{r-1} \dots x_1 x_0)$, которое соответствует Y , осуществляется на основе систем (1) ограничений. Как следует из (1), преобразование выполняется путем отбрасывания подряд стоящих единиц или нулей. Количество исключаемых двоичных разрядов составляет $(n-r)$. Тогда номер i разряда биномиального числа X отличается от номера s разряда равновесной комбинации Y на величину $(n-r)$, т.е. $i = s - n + r$. Поскольку минимальное значение $i_{\min} = 0$, то отсюда $s_{\min} = i_{\min} + n - r = n - r$ и, следовательно, должно быть $s \geq n - r$. Теорема 2 доказана.

Теорема 3. Если в биномиальном числе $X = (x_{r-1} \dots x_1 x_0)$ последняя цифра $x_0 = 0$, то сумма единиц в соответствующей X равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_{n-r} 11 \dots 1)$ с нулевого по $(n-r-1)$ -й разряд

$$\sum_{s=0}^{n-r-1} y_s = n - r.$$

Доказательство. Количество единиц в равновесной комбинации Y

$$k = \sum_{s=0}^{n-1} y_s = \sum_{s=0}^{n-r-1} y_s + \sum_{s=n-r}^{n-1} y_s.$$

Так как из теоремы 2 следует $i = s - n + r$, то, перейдя к значениям x_i соответствующего биномиального числа X , при $s \geq n - r$ получим

$$k = \sum_{s=0}^{n-r-1} y_s + \sum_{i=0}^{r-1} x_i = \sum_{s=0}^{n-r-1} y_s + q.$$

Из системы (1) ограничений для числа X при $x_0 = 0$ находим $q = r - (n - k)$. Подставив в предыдущее выражение значение $q = r - (n - k)$, имеем

$$\sum_{s=0}^{n-r-1} y_s = k - q = k - r + (n - k) = n - r.$$

Теорема 3 доказана.

Следствие 1. Количество нулей в соответствующей $X = (x_{r-1} \dots x_1 x_0)$ равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_{n-r} 11 \dots 1)$ с нулевого по $(n-r-1)$ -й разряд равно нулю.

Теорема 4. Если в биномиальном числе $X = (x_{r-1} \dots x_1 x_0)$ последняя цифра $x_0 = 1$, то сумма единиц в соответствующей X равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_{n-r} 00 \dots 0)$ с нулевого по $(n-r-1)$ -й разряд

$$\sum_{s=0}^{n-r-1} y_s = 0.$$

Доказательство. В отличие от доказательства предыдущей теоремы из системы (1) ограничений при $x_0 = 1$ имеем $q = k$. Следовательно, из последнего выражения доказательства теоремы 3:

$\sum_{s=0}^{n-r-1} y_s = k - q = k - k = 0$. Теорема 4 доказана.

Следствие 2. Количество нулей в соответствующей $X = (x_{r-1} \dots x_i \dots x_1 0)$ равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_{n-r} 11 \dots 1)$ с нулевого по $(n-r-1)$ -й разряд равно $(n-r)$.

Введем следующие обозначения: n_s для переменной, которая представляет собой число разрядов в равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$ от нулевого до разряда s включительно:

$$n_s = s + 1; \quad (10)$$

k_s для переменной, которая представляет собой сумму двоичных разрядов равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$, начиная с нулевого и по разряд s включительно, т.е., другими словами, количество двоичных единиц на кодовом отрезке длины n_s :

$$k_s = \sum_{t=0}^s y_t; \quad (11)$$

m_s для переменной, которая представляет собой количество двоичных нулей равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$, начиная с нулевого и по разряд s включительно:

$$m_s = n_s - k_s. \quad (12)$$

Теорема 5 Сумма единичных значений k_s равновесной комбинации $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$ связана с суммой единичных значений q_{s-n+r} биномиального числа $X = (x_{r-1} \dots x_i \dots x_1 x_0)$ при $n-r \leq s \leq n-1$ следующим соотношением:

$$k_s = k - q_{s-n+r}. \quad (13)$$

Доказательство. Количество k двоичных единиц в равновесной комбинации Y

$$k = \sum_{t=0}^{n-1} y_t = \sum_{t=0}^s y_t + \sum_{t=s+1}^{n-1} y_t.$$

Согласно (11) первое слагаемое выражение есть не что иное как значение k_s . Что касается второго слагаемого, то здесь необходимо учесть, что, во-первых, в соответствии с теоремой 2 для номера s разряда равновесной комбинации Y выполняется $s = i + n - r$ и, следовательно, из условия теоремы $n - r \leq s \leq n - 1$ следует $0 \leq i \leq r - 1$; во-вторых, при $n - r \leq s \leq n - 1$ выполняются равенства $y_{n-r+1} = x_1$, $y_{n-r+2} = x_2$, ..., $y_{n-1} = x_{r-1}$. Принимая во внимание сказанное выше и $x_r = 0$, можно заключить, что второе слагаемое представляется как

$$\sum_{t=s+1}^{n-1} y_t = \sum_{t=s-n+r+1}^{r-1} x_t = \sum_{t=s-n+r+1}^r x_t = q_{s-n+r}.$$

Таким образом, $k = k_s + q_{s-n+r}$ или $k_s = k - q_{s-n+r}$. Теорема 5 доказана.

Теорема 6. Число сочетаний $C_{n-r+i}^{k-q_i}$ нумерационной функции (2) биномиальной системы счисления представляет собой k_s -й элемент $(m_s - 1)$ -й последовательности биномиальных коэффициентов (7):

$$j = k_s, \quad i = m_s - 1, \quad (14)$$

где $n - r \leq s \leq n - 1$.

Доказательство. Используя выражения (9) и (13) из теорем 2 и 5, значение i -го биномиального коэффициента $C_{n-r+i}^{k-q_i}$ нумерационной функции (2) можно записать как

$$C_{n-r+i}^{k-q_i} = C_{n-r+s-n+r}^{k-q_{s-n+r}} = C_s^{k_s}.$$

Чтобы определить значение $C_s^{k_s}$ при помощи последовательностей биномиальных коэффициентов (7), приведем его к виду C_{i+j}^j , воспользовавшись соотношениями (10) и (12):

$$C_s^{k_s} = C_{n_s-1}^{k_s} = C_{m_s-1+k_s}^{k_s}.$$

Как следует из полученного равенства, $j = k_s$, $i = m_s - 1$. Теорема 6 доказана.

Из соотношения (14) видно, что указателем на номер последовательности биномиальных коэффициентов является количество нулей в равновесной комбинации. При этом важно заметить, что новая последовательность рассчитывается на основе предыдущей и, следовательно, расчет чисел сочетаний должен начинаться с нулевой последовательности. Как видно из следствия 2, при переходе от равновесной комбинации Y к биномиальному числу X , для случая, когда разряд $x_0 = 1$, отсчет количества нулей начинается со значения $(n - r)$. Следовательно, возникает необходимость в вычислении неиспользуемых $(n - r)$ последовательностей. Это накладывает дополнительные временные затраты и уменьшает скорость преобразования равновесной комбинации Y в ее порядковый номер X . Для устранения этого недостатка воспользуемся свойством симметрии чисел сочетаний

$$C_{k_s+m_s-1}^{k_s} = C_{k_s+m_s-1}^{m_s-1}. \text{ При этом теперь}$$

$$j = m_s - 1, i = k_s. \quad (15)$$

Таким образом, алгоритм разбивается на две ветви. Первая ветвь соответствует случаю, когда последняя цифра биномиального числа X равна $x_0 = 1$ ($y_0 = 0$). Вычисление значения нумерационной функции (2) осуществляется при помощи суммирования значений $(m_s - 1)$ ячеек k_s последовательностей биномиальных коэффициентов в соответствии с равенствами (15). Вторая ветвь алгоритма соответствует случаю, когда последняя цифра биномиального числа X равна $x_0 = 0$ ($y_0 = 1$). Для этого случая вычисление значения нумерационной функции (2) осуществляется при помощи суммирования значений k_s ячеек $(m_s - 1)$ последовательностей биномиальных коэффициентов в соответствии с равенствами (14).

Количество чисел сочетаний, располагающихся в последовательности биномиальных коэффициентов, для первого случая составляет $(\max m_s - 1) + 1 = (m - 1) + 1 = m$, а для второго – $\max k_s + 1 = k + 1$.

Блок-схема разработанного алгоритма представлена на рис. 2.

Для упрощения записи блок-схемы алгоритма обозначим через M_j переменную со значением биномиального коэффициента, расположенного в j -м элементе текущей i -й последовательности чисел сочетаний.

Рассмотрим работу алгоритма более подробно. Имеется некоторая равновесная комбинация $Y = (y_{n-1} \dots y_s \dots y_1 y_0)$ длиной n бит, содержащая k единиц (вес равновесной комбинации) и m нулей, номер которой необходимо вычислить.

Изначально производится обнуление счетчика единиц k_s , счетчика нулей m_s , счетчика позиции s и ячейки памяти N , в которой будет храниться номер F равновесной комбинации. Также на данном этапе формируется нулевая последовательность чисел сочетаний, содержащая единицы во всех элементах M_j . Счетчик позиции s определяет номер разряда равновесной комбинации Y , с которым работает алгоритм.

На первом этапе преобразования необходимо перейти от равновесной комбинации Y к биномиальному числу X . Из равновесной комбинации считывается s -й бит. Если соседние биты y_s и y_{s-1} равны между собой, то счетчик позиции s увеличивает свое значение на единицу и выполняется новый цикл проверки. В противном случае этап перехода от равновесной комбинации Y к биномиальному числу X считается завершенным. Счетчик позиции s будет указывать на младший разряд x_0 биномиального числа. Также на данном этапе производится подсчет количества k_s единиц и количества m_s нулей для текущего значения s .

На втором этапе производится вычисление номера F исходной равновесной комбинации Y . Как указывалось выше, в зависимости от значения y_0 алгоритм разделяется на две ветви. Первая ветвь соответствует случаю, когда $y_0 = 0$. Если значение y_s -го бита равно

нулю, то инкрементируется счетчик нулей m_s . В противном случае инкрементируется счетчик единиц k_s и вычисляется новая k_s -я последовательность биномиальных коэффициентов.

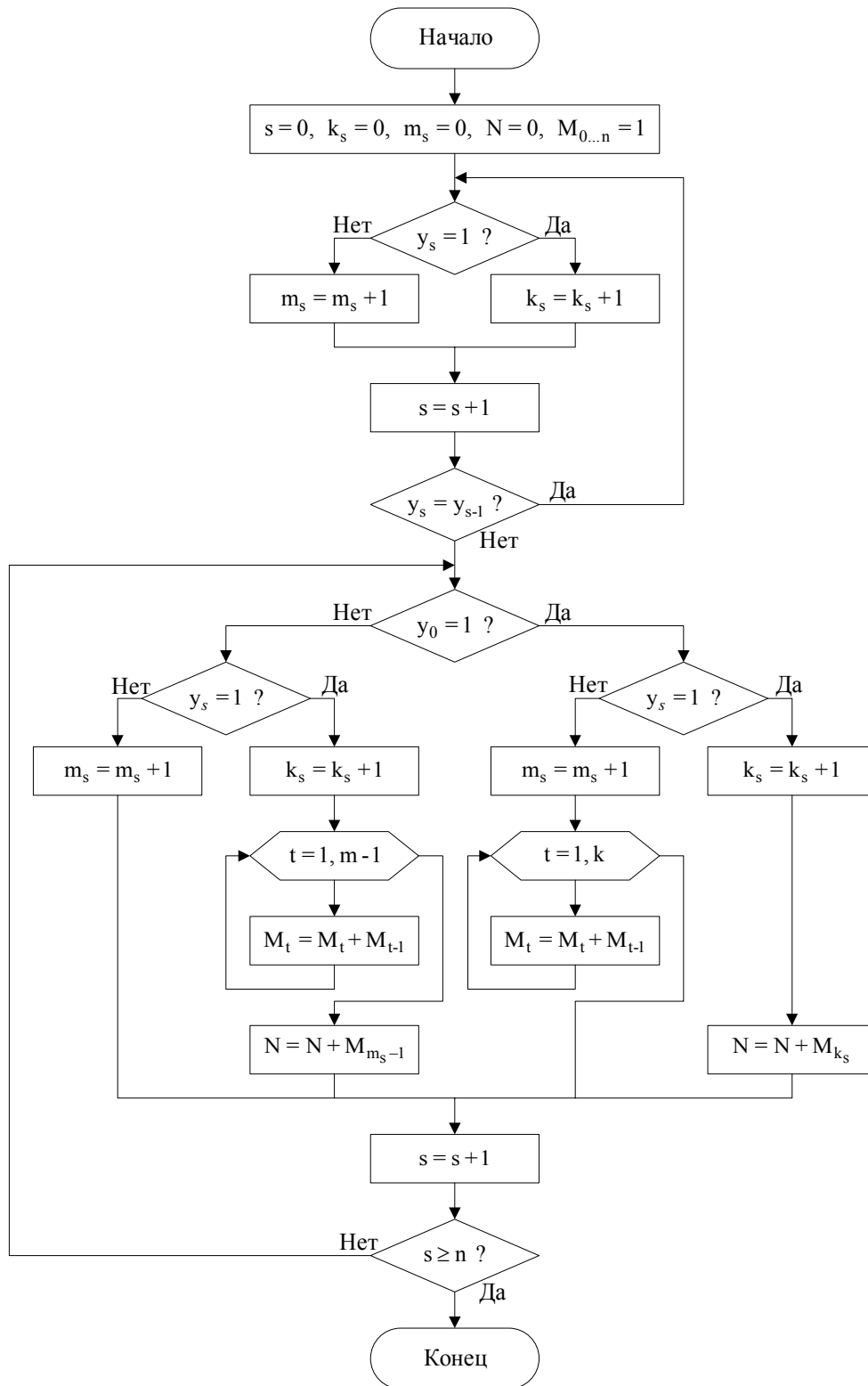


Рис. 2. Блок-схема комбинаторного алгоритма вычисления номера равновесной комбинации

Вычисление новой последовательности производится по циклу в соответствии с выражением (8). После перерасчета последовательности к номеру равновесной комбинации,

который хранится в ячейке памяти N , прибавляется (M_{m_s-1}) -й элемент текущей k_s -й последовательности биномиальных коэффициентов. Вторая ветвь, которая соответствует случаю $y_0 = 1$, выполняется аналогичным образом, с той лишь разницей, что указателем на номер последовательности является счетчик нулей m_s , а указателем на элемент последовательности является счетчик единиц k_s . После выполнения указанных операций счетчик позиции s увеличивается на единицу. Если содержимое счетчика позиции не достигло величины n , то производится выборка следующего y_s -го разряда из равновесной комбинации и цикл вычисления номера повторяется. В противном случае определение номера считается завершенным, а искомое значение хранится в ячейке памяти N .

5. Пример работы алгоритма вычисления номера двоичной комбинации

Согласно методу биномиального нумерационного кодирования вычислить номер для двоичной последовательности 11010100. Параметры заданной комбинации следующие: длина $n = 8$, количество единиц $k = 4$, количество нулей $m = 4$.

Сначала производится обнуление счетчиков $s = 0$, $k_s = 0$, $m_s = 0$ и ячейки памяти $N = 0$, хранящей искомый номер равновесной комбинации. Формируется нулевая последовательность биномиальных коэффициентов, состоящая из четырех элементов: 1 1 1 1.

Считывается нулевой разряд равновесной кодовой комбинации. Поскольку $y_0 = 0$, инкрементируется счетчик нулей $m_s = 1$ и осуществляется переход к следующему разряду с индексом $s = 1$. Поскольку разряды y_0 и y_1 равновесной комбинации равны между собой, повторяются те же действия, что и для нулевого разряда. Значения разрядов y_1 и y_2 уже не равны. Следовательно, двоичное биномиальное число, соответствующее исходной последовательности, получено и преобразование 1 завершено. При этом счетчики позиции, единиц и нулей содержат следующие значения: $s = 2$, $k_s = 0$, $m_s = 2$, а биномиальное число имеет вид 110101.

Поскольку нулевой разряд равновесной комбинации $y_0 = 0$, выполняется первая ветвь алгоритма. Так как $y_2 = 1$, то содержимое счетчика единиц увеличивается на единицу и становится $k_s = 1$. Счетчик нулей при этом сохраняет свое состояние $m_s = 2$. Далее, производится расчет первой последовательности биномиальных коэффициентов, которая теперь будет иметь следующий вид: 1 2 3 4. К содержимому ячейки памяти $N = 0$, хранящей номер равновесной комбинации, прибавляется значение $M_{m_s-1} = M_1$ первого элемента последовательности биномиальных коэффициентов, т.е. $N = N + M_1 = 2$.

Далее рассматриваются остальные разряды биномиального числа и выполняются аналогичные действия до тех пор, пока счетчик позиции s не достигнет значения $n = 8$.

Последовательности значений переменных для каждого цикла алгоритма представлены в табл. 2.

Таблица 2. Пример вычисления номера равновесной комбинации 11010100

s	y_s	k_s	m_s	k_s -я последовательность				Бином коэф.	N
0	0	0	1	1	1	1	1	---	0
1	0	0	2	1	1	1	1	---	0
2	1	1	2	1	2	3	4	2	2
3	0	1	3	1	2	3	4	---	2
4	1	2	3	1	3	6	10	6	8
5	0	2	4	1	3	6	10	---	8
6	1	3	4	1	4	10	20	20	28
7	1	4	4	1	5	15	35	35	63

В результате работы алгоритма был получен номер равновесной комбинации $N_{[11010100]} = 63_{\langle 10 \rangle} = 00111111_{\langle 2 \rangle}$.

6. Выводы

Разработанный комбинаторный алгоритм биномиальной нумерации двоичных комбинаций на основе рекуррентного вычисления последовательностей биномиальных коэффициентов позволяет достигнуть:

- достаточно существенного увеличения быстродействия метода биномиального нумерационного кодирования;
- снижения объема аппаратно-программных затрат при практической реализации биномиальной нумерации двоичных комбинаций.

Достигнутые результаты стали возможными благодаря исключению времязатратных операций умножения и деления над параметрами $k - q_i$, $n - r + i$ чисел сочетаний $C_{n-r+i}^{k-q_i}$, представляющих весовые коэффициенты числовой функции (2), и замене их простыми операциями сложения. Кроме того, снимается проблема переполнения разрядной сетки во время вычисления значений биномиальных коэффициентов, так как исчезает необходимость определения факториалов верхнего и нижнего параметров $C_{n-r+i}^{k-q_i}$. Полученный комбинаторный алгоритм биномиального нумерационного кодирования приобретает особую практическую значимость при больших значениях n и k сжимаемых двоичных последовательностей.

Список литературы: 1. Амелкин В.А. Методы нумерационного кодирования. Новосибирск: Наука, 1986. 155 с. 2. Cover, Thomas M. Enumerative Source Encoding / IEEE Transactions on Information Theory, vol. IT-19, No. 1. Jan. 1973. P. 73-77. 3. Борисенко А.А. Биномиальный счет. Теория и практика: Монография. Сумы: ИТД “Университетская книга”, 2004. 170 с. 4. Кулик И.А., Костель С.В., Скордина Е.М. Использование биномиальных чисел для сжатия бинарных изображений / Вісник СумДУ. 2009. № 2. с. 29-36. 5. Кулик И.А., Скордина Е.М. Метод вычисления биномиальных коэффициентов на основе канонического разложения чисел / Вісник СумДУ. 2008. № 1. С. 158-165. 6. Борисенко А.А. Введение в теорию биномиального счета: Монография. Сумы: ИТД “Университетская книга”, 2004. 89 с.

Поступила в редколлегию 12.11.2009

Кулик Игорь Анатольевич, канд. техн. наук, доцент кафедры электроники и компьютерной техники Сумского государственного университета. Научные интересы: системы сжатия и передачи данных, комбинаторное кодирование. Адрес: Украина, 40007, Сумы, ул. Римско-Корсакова, 2, тел. (0542) 33-55-39. e-mail: kulik@pe.sumdu.edu.ua.

Костель Сергей Викторович, аспирант кафедры электроники и компьютерной техники Сумского государственного университета. Научные интересы: нумерационное кодирование и сжатие информации, программирование. Адрес: Украина, 40007, Сумы, ул. Римско-Корсакова, 2, тел. (0542) 33-55-39. e-mail: eikt@ukr.net.