

ОБ ОДНОЙ КОНСТРУКЦИИ ЛЕГКО ДЕКОДИРУЕМЫХ СУБДЕБРЁЙНОВЫХ МАССИВОВ

Д. А. Макаров^{1,a}, А. Д. Яшунский^{1,2,b}

¹Институт прикладной математики им. М. В. Келдыша РАН,
Миусская пл., 4, 125047 Москва, Россия

²Московский гос. университет им. М. В. Ломоносова,
Ленинские горы, 1, 119991 Москва, Россия

E-mail: ^am8er_ed@mail.ru, ^byashunsky@keldysh.ru

Аннотация. Рассматриваются двумерные обобщения последовательностей де Брёйна — целочисленные массивы, в которых требуется, чтобы все фрагменты заданного размера (окна) были различны. Для таких массивов, называемых субдебрёиновыми, рассматривается сложность задачи декодирования — определения положения в массиве окна с заданным содержимым. Предложена конструкция массивов произвольного размера с произвольными окнами, для которых число различных элементов в массиве по порядку оптимально, а сложность декодирования окон линейна. Библиогр. 16.

Ключевые слова: последовательность де Брёйна, массив де Брёйна, декодирование, сложность.

1. Определения

Массивом A размера $N \times M$ называется индексированный набор элементов $A(i, j) \in \mathbb{N} \cup \{0\}$, где $0 \leq i < N$, $0 \leq j < M$. Фактически массив представляет собой матрицу размера $N \times M$ с целыми неотрицательными элементами, в которой строки и столбцы для удобства нумеруются с нуля. Если для всех i, j выполнено $A(i, j) < c$, то массив A является c -ичным. Значения элементов $A(i, j)$ часто называют *цветами*.

Окном размера $n \times t$, начинающимся в s -й строке и t -м столбце массива A размера $N \times M$, называется массив $[A]_{s,t}^{n \times t}$, элементы которого удовлетворяют равенствам

$$[A]_{s,t}^{n \times t}(i, j) = A(i + s \bmod N, j + t \bmod M)$$

при всех $0 \leq i < n$, $0 \leq j < m$. Здесь и далее предполагается вполне естественное ограничение $n \leq N$, $m \leq M$; значения N , M , n , m , удовлетворяющие указанным неравенствам, будем далее считать допустимыми, говоря о массивах с окнами определённых размеров.

Для массива A размера $N \times M$ окна $[A]_{s,t}^{n \times m}$, у которых s и t удовлетворяют неравенствам $0 \leq s < N - n + 1$ и $0 \leq t < M - m + 1$, будем называть *апериодическими*.

Массив A размера $N \times M$ обладает свойством *уникальности* (апериодических) окон $n \times m$, если для любых пар $(s, t) \neq (s', t')$ таких, что $0 \leq s, s' < N$, $0 \leq t, t' < M$ (соответственно $0 \leq s, s' < N - n + 1$, $0 \leq t, t' < M - m + 1$) выполнено соотношение

$$[A]_{s,t}^{n \times m} \neq [A]_{s',t'}^{n \times m}.$$

Множество массивов $N \times M$ с уникальностью (апериодических) окон $n \times m$ будем обозначать через $\mathcal{U}(N, M, n, m)$ (соответственно $\mathcal{U}'(N, M, n, m)$).

Для массива A размера $N \times M$ через $C(A)$ обозначим число различных цветов в массиве A . Очевидно, что $C(A) \leq NM$. Введём также следующие величины:

$$C_{\mathcal{U}}(N, M, n, m) = \min_{A \in \mathcal{U}(N, M, n, m)} C(A), \quad (1)$$

$$C_{\mathcal{U}'}(N, M, n, m) = \min_{A \in \mathcal{U}'(N, M, n, m)} C(A). \quad (2)$$

Величины $C_{\mathcal{U}}(N, M, n, m)$ ($C_{\mathcal{U}'}(N, M, n, m)$) выражают наименьшее количество цветов, достаточное для построения массива размера $N \times M$ с уникальностью (апериодических) окон $n \times m$.

Несложно заметить, что в соотношениях (1), (2) в действительности минимум достаточно брать по таким массивам A , у которых для всех элементов выполнено $A(i, j) \in \{0, \dots, C(A) - 1\}$. В таких массивах наименьший элемент равен 0, наибольший равен $C(A) - 1$ и все значения от 0 до $C(A) - 1$ встречаются среди элементов массива.

Из неравенства $C(A) \leq NM$ вытекают тривиальные верхние оценки $C_{\mathcal{U}}(N, M, n, m) \leq NM$ и $C_{\mathcal{U}'}(N, M, n, m) \leq NM$.

Нижние оценки, получаемые из комбинаторных соотношений, представляются более содержательными. Кроме того, они точные и достигаются при некоторых значениях параметров. Из уникальности (апериодических) окон вытекают неравенства $NM \leq (C(A))^{nm}$ (соответственно $(N - n + 1)(M - m + 1) \leq (C(A))^{nm}$), из которых следуют оценки

$$C_{\mathcal{U}}(N, M, n, m) \geq (NM)^{\frac{1}{nm}}, \quad (3)$$

$$C_{\mathcal{U}'}(N, M, n, m) \geq ((N - n + 1)(M - m + 1))^{\frac{1}{nm}}. \quad (4)$$

Массив $A \in \mathcal{U}(N, M, n, m)$ (соответственно $A \in \mathcal{U}'(N, M, n, m)$), для которого выполнено $A(i, j) \in \{0, \dots, C(A) - 1\}$ и $C(A) = (NM)^{\frac{1}{nm}}$ (соответственно $C(A) = ((N - n + 1)(M - m + 1))^{\frac{1}{nm}}$), будем называть (*апериодическим*) массивом де Брёйна или *дебрёйновым массивом*, также встречается наименование *совершенный массив*. Все прочие массивы с уникальностью (*апериодических*) окон, для которых выполнено $A(i, j) \in \{0, \dots, C(A) - 1\}$ при всех i и j , будем называть *субдебрёйновыми*.

Дебрёйновы массивы являются двумерным обобщением конструкции, известной как последовательности де Брёйна [4]. С точки зрения данных выше определений s -ичная последовательность де Брёйна с окном длины m является дебрёйновым массивом из множества $\mathcal{U}(1, c^m, 1, m)$.

Уникальность окон в массиве позволяет по заданному содержимому окна вычислять его положение в массиве, т. е. по заданному массиву B , $B = [A]_{s,t}^{n \times m}$, находить значения s и t . Будем называть эту задачу *задачей декодирования*. Сложность её решения определяют как количество операций с элементами массива B (а именно, арифметических действий и операций сравнения), необходимых для нахождения положения левого верхнего элемента окна B в массиве A . Сложность декодирования окна B из массива A будем обозначать через $L(B, A)$.

Пусть $A \in \mathcal{U}(N, M, n, m)$ ($A \in \mathcal{U}'(N, M, n, m)$). Множество всех окон (соответственно всех апериодических окон) размера $n \times m$ в массиве A обозначим через $W(A)$ и введём величину $L(A) = \max_{B \in W(A)} L(B, A)$.

Тривиальное решение задачи декодирования заключается в том, чтобы последовательно сравнить массив B со всеми окнами $n \times m$ в массиве A . Отсюда получаются оценка $L(A) \leq nmNM$ для $A \in \mathcal{U}(N, M, n, m)$ и оценка $L(A) \leq nm(N - n + 1)(M - m + 1)$ для $A \in \mathcal{U}'(N, M, n, m)$. Подобное решение при этом предполагает, что массив A должен храниться в памяти вычисляющего устройства. Такое решение вряд ли можно считать приемлемым с прикладной точки зрения.

Отметим, что, полагая $A(i, j) = iM + j$, $0 \leq i < N$, $0 \leq j < M$, получим массив A с $C(A) = NM$ и $A \in \mathcal{U}(N, M, n, m)$ для любых допустимых значений n и m . При этом декодирование любого окна может быть осуществлено по его единственному элементу и требует константного (не зависящего от N , M , n и m) числа операций. Таким образом, за счёт избыточного использования цветов можно добиться тривиальности решения задачи декодирования.

С прикладной точки зрения представляют интерес массивы, в которых как количество цветов $C(A)$, так и сложность декодирования $L(A)$, по возможности малы. В силу оценок (3) и (4) величина $C(A)$ ограничена снизу и, естественно, оптимальными представляются массивы,

в которых нижняя оценка величины $C(A)$ достигается. Однако для дебрёйновых массивов сложность декодирования оказывается ограниченной снизу. Так как в массиве встречаются всевозможные окна заданного размера, для декодирования потребуется учесть все элементы окна, откуда $L(A) \geq nt$. Таким образом, оптимальность использования цветов априори делает задачу декодирования более трудоёмкой.

2. Обзор результатов

Двумерные обобщения последовательностей де Брёйна предлагались различными авторами. Одной из мотиваций этих обобщений было прикладное значение подобных массивов: в частности, возможность их использования для позиционирования мобильных роботов в подготовленной среде.

Обзор имевшихся по состоянию на 1992 г. результатов в этой области содержится в работе Дж. Бернса и К. Дж. Митчелла [5]. Как отмечается в этом обзоре, прикладное значение субдебрёйновых массивов ничуть не меньше, чем у дебрёйновых. Однако большинство работ было посвящено конструированию именно дебрёйновых массивов, вероятно в силу их комбинаторного интереса. Возможно, единственным исключением являлась работа Дж. Денеса и А. Д. Кидвелла [6], где приведены конструкции именно субдебрёйновых массивов.

В дальнейшем основные исследования касались также различных конструкций дебрёйновых массивов как с произвольными окнами (см. [9, 13, 14]), так и с апериодическими (см., например, [12]). Кроме того, изучался ряд смежных вопросов: в частности, многомерные обобщения дебрёйновых массивов [7], массивы с возможностью исправления ошибок при декодировании окон [2], массивы с окнами в форме креста [3].

В отношении задачи декодирования массивов долгое время исследования были направлены на построение для существующих дебрёйновых массивов алгоритмов декодирования, отличных от тривиальных. В частности, такие алгоритмы для некоторых массивов приведены в [10, 11, 15]. В работе Дж. Тулиани [16] были предложены специальным образом сконструированные последовательности де Брёйна с окном размера $1 \times t$, декодируемые за $O(m \log t)$ операций, с использованием $O(1)$ памяти.

В недавней работе В. Хоран и Б. Стивенса [8] окна в форме креста из [3] обобщаются до более широкого класса «нестандартных» окон — некоторых специальных подмножеств прямоугольника $n \times t$ (причём собственно прямоугольное окно $n \times t$ в этот класс не попадает). В [8] приводятся конструкции массивов, в которых сложность декодирования рассматриваемых видов окон составляет $O(n^2 + t^2)$. Отметим, что при $n = t$ это совпадает по порядку с нижней оценкой сложности декодирования окна в дебрёйновом массиве.

Настоящая работа является прямым развитием идей, предложенных Д. А. Макаровым в [1]. В ней описана конструкция субдебрёйновых массивов, для которых величина $L(A)$ по порядку равна $O(nm)$, а величина $C(A)$ есть $O((NM)^{\frac{1}{nm}})$, т. е. оптимальна по порядку.

3. Канонические массивы

Рассмотрим некоторые специальные конструкции субдебрёйновых массивов с заданным количеством цветов c . Зафиксируем некоторое число d , $\frac{c}{2} \leq d < c$, и будем называть цвета $d, \dots, c-1$ *маркерными*, а прочие цвета — *обычными*.

Пусть Q — c -ичный массив размера $n \times m$, для которого выполнено $Q(i, j) < d$ при $(i, j) \neq (n-1, m-1)$ и $Q(n-1, m-1) \geq d$. Тогда Q будем называть *каноническим окном*. Массиву Q сопоставим число ξ , вычисляемое следующим образом:

$$\xi = -d^{nm} + \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} d^{im+j} Q(i, j). \quad (5)$$

Фактически ξ есть число, выражаемое содержимым массива Q , если его рассматривать как d -ичную запись числа от младшего разряда к старшему (слева направо и сверху вниз), в которой в качестве самого старшего разряда берётся значение $Q(n-1, m-1) - d$ (из условия $d \geq \frac{c}{2}$ вытекает, что $Q(n-1, m-1) - d < d$).

Положим $\sigma = (c-d)d^{nm-1}$. Несложно заметить, что между каноническими окнами и числами ξ , удовлетворяющими $0 \leq \xi < \sigma$, имеется взаимно однозначное соответствие. Как следствие, при заданных параметрах n, m, c, d каждому числу ξ , $0 \leq \xi < \sigma$, соответствует единственное каноническое окно, которое будем обозначать через $Q_{\xi}^{n,m,c,d}$.

Ниже будет представлена конструкция массивов, основанная на использовании канонических окон $Q_{\xi}^{n,m,c,d}$. Мы будем формировать из них прямоугольные массивы путем «склейки». Очевидно, что условие уникальности окон ограничивает площадь таких массивов сверху, так как ни одно из канонических окон не может быть использовано целиком дважды. При этом соотношение числа строк и столбцов в формируемом массиве, вообще говоря, может меняться за счёт изменения числа канонических окон, используемых в каждой строке массива. В формальном описании конструкции, представленном далее, количество канонических окон, целиком используемых в каждой строке массива, обозначено через μ , а в каждом столбце массива — через ν .

Для таких массивов будет затем показана однозначность декодирования произвольных апериодических окон и, как следствие, принадлежность классу $\mathcal{U}'(N, M, n, m)$ для некоторых значений N и M . В разд. 4

субдебрёйновы массивы произвольного размера будут формироваться путём выделения подходящих подмассивов в массивах построенной серии.

Далее считаем, что параметры $n, m, c, d \in \mathbb{N}$ фиксированы и d удовлетворяет условию $\frac{c}{2} \leq d < c$. Выберем некоторые $\nu, \mu \in \mathbb{N}$, удовлетворяющие неравенству $\nu\mu \leq \sigma$, и определим массив P размера $(\nu+1)n \times (\mu+1)m$ так, чтобы при всех i, j выполнялось соотношение

$$[P]_{in,jm}^{n \times m} = Q_{(i\mu+j) \bmod \sigma}^{n,m,c,d}. \quad (6)$$

Несложно проверить, что такое определение массива P корректно. Обозначая для краткости $Q_{\xi}^{n,m,c,d}$ через Q_{ξ} , массив P можно записать в следующем виде:

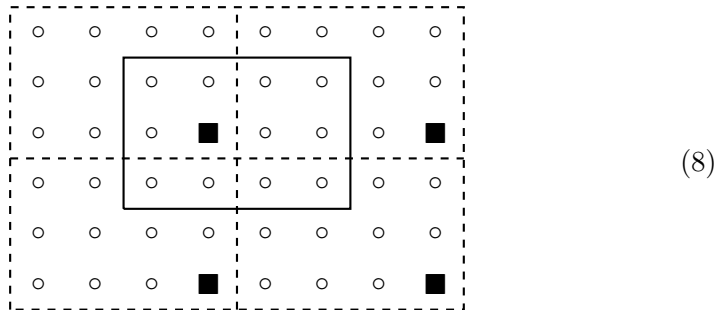
$$P = \begin{bmatrix} Q_0 & Q_1 & \dots & Q_{\mu} \\ Q_{\mu} & Q_{\mu+1} & \dots & Q_{2\mu} \\ \dots & \dots & \dots & \dots \\ Q_{(\nu-2)\mu} & Q_{(\nu-2)\mu+1} & \dots & Q_{(\nu-1)\mu} \\ Q_{(\nu-1)\mu} & Q_{(\nu-1)\mu+1} & \dots & Q_{\nu\mu \bmod \sigma} \\ Q_{\nu\mu \bmod \sigma} & Q_{(\nu\mu+1) \bmod \sigma} & \dots & Q_{(\nu+1)\mu \bmod \sigma} \end{bmatrix}. \quad (7)$$

Здесь и далее подобную запись (массив из массивов) будем рассматривать как «склеивание» меньших массивов в единый массив.

Лемма 1. Любое окно размера $n \times t$ в массиве P содержит маркерный элемент и при том единственный.

Доказательство. Массив P фактически составлен из канонических окон размера $n \times t$, в каждом из которых имеется единственный маркерный элемент.

Рассмотрим произвольное окно $n \times t$ в массиве P . Ниже схематически представлено положение окна в массиве: символ \blacksquare обозначает маркерные элементы, символ \circ — обычные, штриховой линией показаны границы канонических окон, сплошной — границы окна.



Окна, не являющиеся апериодическими, также могут считаться расположенными указанным выше образом относительно канонических: для этого их надо рассматривать как апериодические окна в массиве $\begin{bmatrix} P & P \\ P & P \end{bmatrix}$, что соответствует определению произвольного окна в массиве P .

Несложно видеть, что в любом окне размера $n \times m$ также будет присутствовать единственный маркерный элемент. Его расположение в окне будет показывать, насколько рассматриваемое окно смещено относительно канонических. Лемма 1 доказана.

Замечание 1. Нахождение местоположения маркерного элемента в заданном окне $n \times m$ требует не более nm операций сравнения.

Докажем одно вспомогательное утверждение о числах в d -ичной системе счисления. Напомним, что число α единственным образом представляется в виде $\alpha = \sum_{l \geq 0} d^l \alpha_l$, где α_l — разряды числа α в d -ичной системе счисления, $0 \leq \alpha_l < d$.

Лемма 2. Зафиксируем натуральные параметры r, d, d' ($0 < d' \leq d$) и положим $\delta = d'd^{r-1}$. Пусть некоторые числа $0 \leq \xi, \zeta_1, \dots, \zeta_s, \theta_1, \dots, \theta_s < \delta$ связаны соотношениями $\theta_t = (\xi + \zeta_t) \bmod \delta$, $t = 1, \dots, s$, и d -ичное разложение чисел θ_t имеет вид $\theta_t = \sum_{l=0}^{r-1} d^l \theta_{t,l}$ для $t = 1, \dots, s$.

Пусть заданы числа ζ_1, \dots, ζ_s , а также такой набор чисел $\alpha_0, \dots, \alpha_{r-1}$, что для каждого $l = 0, \dots, r-1$ найдётся номер t_l , для которого выполнено $\theta_{t_l,l} = \alpha_l$. Тогда число ξ определено однозначно.

ДОКАЗАТЕЛЬСТВО. Рассмотрим сложение чисел $\xi + \zeta_t$, $t = 1, \dots, s$, в d -ичной системе счисления. Отметим, что вычисление указанных сумм по модулю δ влияет только на старший разряд в каждой сумме — фактически он вычисляется по модулю d' . Пусть d -ичное разложение чисел ζ_t имеет вид $\zeta_t = \sum_{l=0}^{r-1} d^l \zeta_{t,l}$, $t = 1, \dots, s$. Также для каждого $t = 1, \dots, s$ и $l = 0, \dots, r-1$ обозначим через $\eta_{t,l}$ перенос, возникающий из l -го в $(l+1)$ -й разряд при сложении ξ и ζ_t .

Покажем, что в условиях леммы все разряды в d -ичном представлении числа $\xi = \sum_{l=0}^{r-1} d^l \xi_l$ определены однозначно. При $l = 0$ по условию леммы найдётся такое t_0 , что задано значение $\theta_{t_0,0} = (\xi_0 + \zeta_{t_0,0}) \bmod d = \alpha_0$. Поскольку $\zeta_{t_0,0}$ также известно, можно вычислить

$$\xi_0 = (\alpha_0 - \zeta_{t_0,0}) \bmod d.$$

Далее, используя полученное ξ_0 и известные значения $\zeta_{1,0}, \dots, \zeta_{s,0}$, можно также вычислить все переносы $\eta_{t,0}$ при $t = 1, \dots, s$.

Рассмотрим теперь последовательно $l = 1, \dots, r - 2$. Для каждого значения l найдётся такое t_l , что задано $\theta_{t_l, l} = \alpha_l$. Поскольку выполнено равенство $\alpha_l = \theta_{t_l, l} = (\xi_l + \zeta_{t_l, l} + \eta_{t_l, l-1}) \bmod d$, а значение $\eta_{t_l, l-1}$ найдено на предыдущем шаге, получаем

$$\xi_l = (\alpha_l - \zeta_{t_l, l} - \eta_{t_l, l-1}) \bmod d.$$

Так же, как и в случае $l = 0$, поскольку известны все значения $\zeta_{1, l}, \dots, \zeta_{s, l}$, легко вычисляются все переносы $\eta_{t, l}$ при $t = 1, \dots, s$.

Случай $l = r - 1$ отличается от рассмотренных выше только тем, что вычисления проводятся по модулю d' :

$$\xi_{r-1} = (\alpha_{r-1} - \zeta_{t_{r-1}, r-1} - \eta_{t_{r-1}, r-2}) \bmod d'.$$

Итак, все r разрядов числа ξ в d -ичной системе счисления определены однозначно. Лемма 2 доказана.

Замечание 2. Несложно видеть, что в условиях леммы определение каждого из разрядов числа ξ требует $O(1)$ арифметических операций. Как следствие, вычисление ξ осуществляется за $O(r)$ операций.

Определим два массива R и R' , зависящие от параметров n, m, c, d, ν, μ , положив $R = [P]_{0,0}^{\nu n \times \mu m}$, $R' = [P]_{0,0}^{(\nu+1)n-1 \times (\mu+1)m-1}$.

Лемма 3. Для массива R' выполнены соотношения

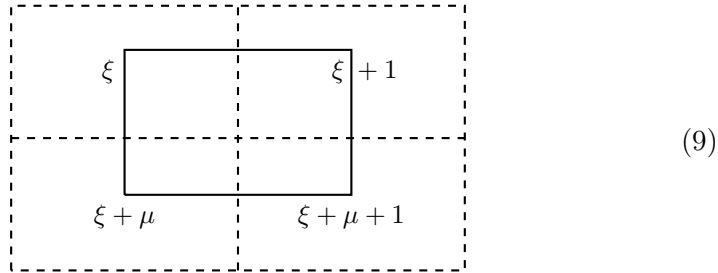
$$R' \in \mathcal{U}'((\nu+1)n-1, (\mu+1)m-1, n, m), \quad L(R') = O(nm).$$

ДОКАЗАТЕЛЬСТВО. Покажем, что произвольное апериодическое окно размера $n \times m$ в массиве R' однозначно декодируется, отсюда будет следовать принадлежность $R' \in \mathcal{U}'((\nu+1)n-1, (\mu+1)m-1, n, m)$.

Пусть задано произвольное апериодическое окно B в массиве R' . По лемме 1 в этом окне имеется единственный маркерный элемент. Если он расположен в правом нижнем углу окна, то окно B совпадает с одним из канонических окон. Отметим, что в силу определения массива R' каждое из канонических окон встречается в нём целиком ровно один раз: как видно из формулы (7), канонические окна $Q_\mu, Q_{2\mu}, Q_{3\mu}, \dots$ имеют фрагменты, которые входят дважды, однако если апериодическое окно B в массиве R' в точности совпало с некоторым каноническим, то его положение определяется однозначно. Для его нахождения достаточно вычислить значение ξ для окна B по формуле (5). Строка и столбец, в которых находится левый верхний элемент окна B , выражаются через ξ как $n \lfloor \frac{\xi}{\mu} \rfloor$ и $m(\xi \bmod \mu)$.

Пусть окно B не совпадает ни с одним из канонических. Тогда, как следует из леммы 1, в нём имеется единственный маркерный элемент и его положение в окне B позволяет понять, насколько окно B смещено относительно границ канонических окон (см. (8)). При этом, поскольку

окно B не совпадает ни с одним из канонических, в него будут попадать некоторые части из не более чем четырёх канонических окон (не ограничивая общности, можно считать, что в рассматриваемом окне имеются фрагменты в точности четырёх канонических окон, при этом некоторые из фрагментов могут быть пустыми). Как изображено ниже, окно составлено из фрагментов канонических окон с номерами ξ , $\xi + 1$, $\xi + \mu$, $\xi + \mu + 1$ для некоторого значения ξ .



Если рассмотреть запись каждого из четырёх чисел ξ , $\xi + 1$, $\xi + \mu$, $\xi + \mu + 1$ в d -ичной системе счисления, то каждое из них будет содержать nm разрядов и для каждого номера $l = 0, \dots, nm - 1$ в окне B будет содержаться l -й разряд какого-то из этих четырёх чисел (с точностью до того, что для $(nm - 1)$ -го разряда надо из соответствующего элемента в окне B вычесть d). Полагая $r = nm$ и $d' = c - d$, $s = 4$, $\zeta_1 = 0$, $\zeta_2 = 1$, $\zeta_3 = \mu$, $\zeta_4 = \mu + 1$, мы попадаем в условия леммы 2, из которой следует, что по содержимому окна B число ξ определяется однозначно.

Зная ξ , а также смещение окна B относительно канонического окна Q_ξ , легко установить положение верхнего левого элемента окна B в массиве R' . Тем самым установлена однозначность декодирования.

Оценка $L(R') = O(nm)$ легко вытекает из замечаний 1 и 2, а также приведённых выше рассуждений о способе декодирования окна. Лемма 3 доказана.

Процедуру декодирования, лежащую в основе доказательства леммы 3, можно представить в виде следующего алгоритма на псевдокоде.

```

procedure DECODE( $B, n, m, c, d, \mu$ )
   $i \leftarrow 0, j \leftarrow 0$ 
  while  $B(i, j) < d$  do
     $j \leftarrow j + 1$ 
    if  $j = m$  then  $i \leftarrow i + 1, j \leftarrow 0$  end if
  end while
   $x \leftarrow j, y \leftarrow i$   ▷ Маркер найден, ищем номер канонического окна  $\xi$ 

```

```

 $\xi \leftarrow 0, \eta_0 \leftarrow 0, \eta_1 \leftarrow 0, \eta_\mu \leftarrow 0, \eta_{\mu+1} \leftarrow 0, l \leftarrow 0, e \leftarrow 1$ 
for  $i \leftarrow 0, \dots, n-1$  do
  for  $j \leftarrow 0, \dots, m-1$  do
     $\varsigma \leftarrow \mu + 1$   $\triangleright$  Меняем  $\varsigma$  так, чтобы  $\xi + \varsigma$  было номером
    if  $j + 1 + x \geq m$  then  $\varsigma \leftarrow \varsigma - 1$  end if  $\triangleright$  канонического
    if  $i + 1 + y \geq n$  then  $\varsigma \leftarrow \varsigma - \mu$  end if  $\triangleright$  окна  $l$ -го разряда.
     $\varsigma_l \leftarrow \lfloor \varsigma / e \rfloor \bmod d$   $\triangleright$   $l$ -й разряд числа  $\varsigma$ 
     $\alpha \leftarrow B(i + 1 + y \bmod n, j + 1 + x \bmod m) \bmod d$ 
    if  $i = n - 1, j = m - 1$  then  $d \leftarrow c - d$  end if
     $\xi_l \leftarrow (\alpha - \varsigma_l - \eta_\varsigma) \bmod d, \xi \leftarrow \xi + e \cdot \xi_l$   $\triangleright$  Вычислен  $l$ -й разряд
    for  $\varsigma \in \{1, \mu, \mu + 1\}$  do  $\triangleright$  Обновление переносов
       $\eta_\varsigma \leftarrow \lfloor (\xi_l + \eta_\varsigma + \lfloor \varsigma / e \rfloor \bmod d) / d \rfloor$ 
    end for
     $l \leftarrow l + 1, e \leftarrow e \cdot d$   $\triangleright$   $l$  — номер следующего разряда,  $e = d^l$ 
  end for
end for
 $x \leftarrow m(\xi \bmod \mu) + (m - 1 - x), y \leftarrow n(\lfloor \xi / \mu \rfloor) + (n - 1 - y)$ 
end procedure  $\triangleright$  В  $x$  и  $y$  координаты верхнего левого угла окна  $B$ 

```

Из алгоритма видно, что требуется не более двух обращений к значению каждого из элементов $B(i, j)$ массива B и помимо памяти под собственно массив B требуется лишь ограниченный (не зависящий от N, M, n, m) объём памяти под все вспомогательные переменные.

С прикладной точки зрения именно случай аperiодических окон представляется наиболее интересным, однако для полноты картины покажем, что массив P может служить основой для построения массива с уникальностью произвольных окон. Для массива R принадлежность к множеству $\mathcal{U}(\nu n, \mu m, n, m)$ достаточно просто доказать, если между параметрами n, m, c, d, ν, μ имеются дополнительные соотношения.

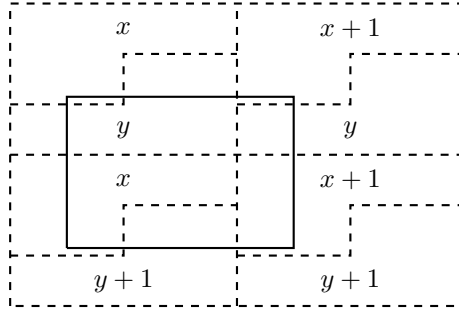
Лемма 4. Пусть существует такое $h \in \mathbb{N}$, что выполнены соотношения $d^h = \mu$ и $(c - d)d^{nm-h-1} = \nu$. Тогда для массива R выполнено $R \in \mathcal{U}(\nu n, \mu m, n, m)$ и $L(R) = O(nm)$.

ДОКАЗАТЕЛЬСТВО. Как и в лемме 3, покажем, что произвольное окно B в массиве R однозначно декодируется за $O(nm)$ операций: из этого сразу же будет следовать уникальность окон. Процедура декодирования окон во многом похожа на соответствующую процедуру из леммы 3, однако окна, не являющиеся аperiодическими, фактически берутся в подмассивах вида $\begin{bmatrix} Q_{i\mu+\mu-1} & Q_{i\mu} \\ Q_{(i+1)\mu+\mu-1} & Q_{(i+1)\mu} \end{bmatrix}$, где расположение канонических окон не соответствует представлению (9) и их декодирование не может

осуществляться описанным ранее образом. Вместе с тем условия $d^h = \mu$ и $(c - d)d^{nm-h-1} = \nu$ позволяют предложить для элементов канонических окон дополнительную интерпретацию, помогающую декодировать произвольные (а не только апериодические) окна.

С каноническим окном Q_ξ можно связать два числа: $x = \xi \bmod \mu$ и $y = \lfloor \xi/\mu \rfloor$. Фактически это номера столбца и строки, в которых стоит Q_ξ в представлении (7). Несложно заметить, что младшие h разрядов в каноническом окне являются записью числа x , а старшие $nm - h$ разрядов — записью числа y в d -ичной системе счисления.

Тогда в смежных канонических окнах будут записаны разряды d -ичной записи чисел x , $(x + 1) \bmod \mu$, y , $(y + 1) \bmod \nu$, часть из которых попадёт в произвольное окно, как изображено ниже: штриховой линией показаны границы канонических окон и их фрагментов, сплошной — границы произвольного окна.



Рассматривая произвольное окно размера $n \times t$ в массиве R , можем применить лемму 2 к двум h -разрядным числам x и $x + 1$ в d -ичной системе счисления с $d' = d$, а затем ещё раз применить лемму 2 к двум $(nm - h)$ -разрядным числам y и $y + 1$ в d -ичной системе счисления для $d' = c - d$. Таким образом, значения x и y , а значит, и положение окна, определены однозначно. Несложно видеть, что для декодирования вновь потребуется $O(nm)$ операций. В силу однозначности декодирования $R \in \mathcal{U}(\nu n, \mu t, n, t)$. Лемма 4 доказана.

4. Массивы, оптимальные по порядку $C(A)$

Покажем, что предложенная выше конструкция субдебрёйновых массивов позволяет в зависимости от параметров N , M , n , t строить массивы A с уникальностью апериодических окон, у которых количество используемых цветов $C(A)$ при фиксированном значении nm и стремлении $N, M \rightarrow \infty$ по порядку совпадает с нижней оценкой величины $C(A)$.

Пусть далее $w = nm$ фиксировано. Для заданных N и M положим $\nu = \lfloor \frac{N}{n} \rfloor$, $\mu = \lfloor \frac{M}{m} \rfloor$. Кроме того, пусть $d' = \lceil \frac{1}{w-1}(\nu\mu(w-1))^{\frac{1}{w}} \rceil$, $c = wd'$ и $d = c - d'$.

Неравенство $d \geq c/2$ равносильно $d' \leq c/2$, т. е. $2d' \leq c$, которое в силу $c = wd'$ очевидно выполнено для $w \geq 2$ (тем самым предлагаемая конструкция не может быть использована для массивов с окном 1×1 , однако для такого окна массивы строятся тривиально).

По выбранным так параметрам n, m, c, d, ν, μ построим массив P , удовлетворяющий соотношениям (6). Такое построение возможно, если выполнено неравенство $\nu\mu \leq \sigma = (c - d)d^{nm-1}$. Проверим, что это неравенство действительно выполняется:

$$\begin{aligned} \sigma &= (c - d)d^{nm-1} = d'(c - d')^{w-1} = d'(wd' - d')^{w-1} \\ &= d'^w(w-1)^{w-1} = (w-1)^{w-1} \left\lceil \frac{1}{w-1}(\nu\mu(w-1))^{\frac{1}{w}} \right\rceil^w \\ &\geq (w-1)^{w-1} \left(\frac{1}{w-1}(\nu\mu(w-1))^{\frac{1}{w}} \right)^w = \frac{(w-1)^{w-1}}{(w-1)^w} \nu\mu(w-1) = \nu\mu. \end{aligned}$$

В силу выбора значений ν и μ очевидно выполнено $N < (\nu + 1)n$ и $M < (\mu + 1)m$, поэтому массив $R' = [P]_{0,0}^{N \times M}$ размера $N \times M$ будет субдебрёиновым с окном $n \times m$. Очевидно, что $C(R') = c$, поэтому для $C_{\mathcal{U}'}(N, M, n, m)$ имеет место следующая оценка:

$$\begin{aligned} C_{\mathcal{U}'}(N, M, n, m) &\leq c = wd' = w \left\lceil \frac{1}{w-1}(\nu\mu(w-1))^{\frac{1}{w}} \right\rceil \\ &= nm \left\lceil \frac{1}{nm-1} \left(\left\lfloor \frac{N}{n} \right\rfloor \left\lfloor \frac{M}{m} \right\rfloor (nm-1) \right)^{\frac{1}{nm}} \right\rceil. \end{aligned}$$

Покажем, что эта верхняя оценка при фиксированных n, m и растущих значениях N, M по порядку совпадает с нижней оценкой, приведённой в (4). Рассмотрим отношение верхней оценки к нижней при фиксированных n, m и $N, M \rightarrow \infty$ (символом \sim обозначаем асимптотическое равенство):

$$\begin{aligned} \frac{nm \left\lceil \frac{1}{nm-1} \left(\left\lfloor \frac{N}{n} \right\rfloor \left\lfloor \frac{M}{m} \right\rfloor (nm-1) \right)^{\frac{1}{nm}} \right\rceil}{((N-n+1)(M-m+1))^{\frac{1}{nm}}} &\sim \frac{\frac{nm}{nm-1} \left(\frac{NM}{nm} (nm-1) \right)^{\frac{1}{nm}}}{(NM)^{\frac{1}{nm}}} \\ &= \frac{nm}{nm-1} \left(\frac{nm-1}{nm} \right)^{\frac{1}{nm}} = \left(\frac{nm-1}{nm} \right)^{\frac{1}{nm}-1} = \left(1 - \frac{1}{nm} \right)^{\frac{1}{nm}-1}. \end{aligned}$$

Таким образом, для каждого фиксированного $w = nm$ отношение верхней оценки величины $C_{\mathcal{U}'}(N, M, n, m)$ к её нижней оценке асимптотически равно $(1 - \frac{1}{w})^{\frac{1}{w}-1}$. Отметим, что эта величина монотонно убывает к 1 с ростом w , и при всех натуральных $w \geq 2$ она оценивается сверху своим значением при $w = 2$, т. е. имеет место неравенство $(1 - \frac{1}{w})^{\frac{1}{w}-1} \leq \sqrt{2}$.

Итак, установлено асимптотическое (при $N, M \rightarrow \infty$) неравенство

$$C_{\mathcal{U}'}(N, M, n, m) \lesssim \left(1 - \frac{1}{nm}\right)^{\frac{1}{nm}-1} ((N - n + 1)(M - m + 1))^{\frac{1}{nm}}.$$

В силу уже использовавшегося выше асимптотического равенства

$$(N - n + 1)(M - m + 1)^{\frac{1}{nm}} \sim (NM)^{\frac{1}{nm}}$$

доказана

Теорема 1. При фиксированных n, m и $N, M \rightarrow \infty$ имеет место соотношение

$$(NM)^{\frac{1}{nm}} \lesssim C_{\mathcal{U}'}(N, M, n, m) \lesssim \left(1 - \frac{1}{nm}\right)^{\frac{1}{nm}-1} (NM)^{\frac{1}{nm}}$$

и, как следствие, $C_{\mathcal{U}'}(N, M, n, m) = O((NM)^{\frac{1}{nm}})$.

Выбор параметров при доказательстве теоремы фактически даёт способ построения легко декодируемого субдебрёйнова массива при заданных величинах N, M, n, m : по ним следует вычислить ν и μ , затем значение d' , а по нему — значения s и d , которые затем использовать для построения канонического массива. Из доказанной теоремы следует, что при таком построении для достаточно больших массивов количество используемых цветов не более чем в $\sqrt{2}$ раз превышает минимально возможное значение. При этом для декодирования окна размера $n \times m$ в построенном массиве потребуется лишь $O(nm)$ операций. Таким образом доказана

Теорема 2. Для любых допустимых значений N, M, n, m найдётся такой массив $R'(N, M, n, m) \in \mathcal{U}'(N, M, n, m)$, что выполнено равенство

$$L(R'(N, M, n, m)) = O(nm),$$

а кроме того при фиксированных n, m и $N, M \rightarrow \infty$

$$C(R'(N, M, n, m)) = O(C_{\mathcal{U}'}(N, M, n, m)).$$

5. Заключение

Полученные результаты имеют очевидную практическую значимость: предложенные алгоритмы построения и декодирования субдебрёиновых массивов могут непосредственно использоваться в прикладных задачах, например, для позиционирования мобильных роботов в специально подготовленной среде. С теоретической точки зрения представляет несомненный интерес вопрос о том, как связаны между собой количество используемых цветов и сложность декодирования, в частности, возможно ли декодирование с линейной сложностью при оптимальном числе используемых цветов? Исследование этой зависимости может стать предметом дальнейших исследований.

А. Д. Яшунский выражает благодарность В. Д. Яшунскому, который привлек внимание автора к задаче построения дебрёиновых массивов.

ЛИТЕРАТУРА

1. **Макаров Д. А.** Построение легко декодируемых субдебрёиновых матриц с окном 2×2 // Материалы X молодёж. науч. школы по дискрет. математике и её прил. (Москва, 5–11 октября 2015 г.) ИПМ им. М. В. Келдыша, 2015. С. 47–50. <http://keldysh.ru/dmschool/datastore/media/sbornikX.pdf>
2. **Berkowitz R., Kopparty S.** Robust positioning patterns // Proc. 27th Annu. ACM-SIAM Symp. Discrete Algorithms SODA'16 (Arlington, VA, Jan. 10–12, 2016). Philadelphia, PA: SIAM, 2016. P. 1937–1951.
3. **Bruckstein A. M., Etzion T., Giryes R., Gordon N., Holt R. J., Shuldiner D.** Simple and robust binary self-location patterns // IEEE Trans. Inf. Theory. 2012. Vol. 58, No. 7. P. 4884–4889.
4. **De Bruijn N. G.** A combinatorial problem // Proc. Nederl. Akad. Wet. 1946. Vol. 49, No. 7. P. 758–764.
5. **Burns J., Mitchell C. J.** Coding schemes for two-dimensional position sensing. Res. Rep. HPL-92-19, Bristol: HP Lab., 1992. <http://www.hpl.hp.com/techreports/92/HPL-92-19.pdf>
6. **Dénes J., Keedwell A. D.** A new construction of two-dimensional arrays with the window property // IEEE Trans. Inf. Theory. 1990. Vol. 36, No. 4. P. 873–876.
7. **Etzion T.** Sequence folding, lattice tiling, and multidimensional coding // arXiv:0911.1745. <http://arxiv.org/abs/0911.1745v1> (2009)
8. **Horan V., Stevens B.** Locating patterns in the de Bruijn torus // Discrete Math. 2016. Vol. 339, No. 4. P. 1274–1282. doi: 10.1016/j.disc.2015.11.015
9. **Hurlbert G., Isaak G.** On the de Bruijn torus problem // J. Comb. Theory, Ser. A. 1993. Vol. 64, No. 1. P. 50–62.
10. **Lloyd S. A., Burns J.** Finding the position of a subarray in a pseudo-random array // Cryptography and Coding III. Oxford: Clarendon Press, 1993.
11. **Mitchell C. J., Paterson K. G.** Decoding perfect maps // Des. Codes Cryptogr. 1994. Vol. 4, No. 1. P. 11–30.

12. **Mitchell C. J.** Aperiodic and semi-periodic perfect maps // IEEE Trans. Inf. Theory. 1995. Vol. 41, No. 1. P. 88–95.
13. **Paterson K. G.** New classes of perfect maps I // J. Comb. Theory, Ser. A. 1996. Vol. 73, No. 2. P. 302–334.
14. **Paterson K. G.** New classes of perfect maps II // J. Comb. Theory, Ser. A. 1996. Vol. 73, No. 2. P. 335–345.
15. **Shiu W. C.** Decoding de Bruijn arrays constructed by FFMS method // Ars Comb. 1997. Vol. 47. P. 33–48.
16. **Tuliani J.** De Bruijn sequences with efficient decoding algorithms // Discrete Math. 2001. Vol. 226, No. 1–3. P. 313–336.

Макаров Дмитрий Александрович
Яшунский Алексей Дмитриевич

Статья поступила
30 октября 2018 г.
После доработки —
14 февраля 2019 г.
Принята к публикации
27 февраля 2019 г.

ON A CONSTRUCTION OF EASILY DECODABLE SUB-DE BRUIJN ARRAYS

D. A. Makarov^{1,a} and A. D. Yashunsky^{1,2,b}

¹Keldysh Institute of Applied Mathematics,
4 Miusskaya Square, 125047 Moscow, Russia

²Lomonosov Moscow State University,
1 Leninskie gory, 119991 Moscow, Russia

E-mail: ^am8er_ed@mail.ru, ^byashunsky@keldysh.ru

Abstract. We consider a two-dimensional generalization of de Bruijn sequences; i.e., integer-valued arrays whose all fragments of a fixed size (windows) are different. For these arrays, dubbed sub-de Bruijn, we consider the complexity of decoding; i.e., the determination of a position of a window with given content in an array. We propose a construction of arrays of arbitrary size with arbitrary windows where the number of different elements in the array is of an optimal order and the complexity of decoding a window is linear. Bibliogr. 16.

Keywords: de Bruijn sequence, de Bruijn array, decoding, complexity.

REFERENCES

1. D. A. Makarov, Construction of easily decodable sub-de Bruijn matrices with a 2×2 window, in *Materialy X molodyozhnoi nauchnoi shkoly po diskretnoi matematike i eyo prilozheniyam* (Proc. 10th Young Scientists' School on Discrete Mathematics and Its Applications) *Moscow, Russia, Oct. 5–11, 2015*, pp. 47–50, Keldysh Inst. Appl. Math., Moscow, 2015 [Russian]. Available at <http://keldysh.ru/dmschool/datastore/media/sbornikX.pdf> (accessed Feb. 25, 2019).
2. R. Berkowitz and S. Kopparty, Robust positioning patterns, in *Proc. 27th Ann. ACM-SIAM Symp. Discrete Algorithms, Arlington, VA, USA, Jan. 10–12, 2016*, pp. 1937–1951, SIAM, Philadelphia, PA, 2016.
3. A. M. Bruckstein, T. Etzion, R. Giryes, N. Gordon, R. J. Holt, and D. Shuldiner, Simple and robust binary self-location patterns, *IEEE Trans. Inf. Theory*, **58**, No. 7, 4884–4889, 2012.
4. N. G de Bruijn, A combinatorial problem, *Proc. Nederl. Akad. Wet.*, **49**, No. 7, 758–764, 1946.

5. **J. Burns** and **C. J. Mitchell**, Coding schemes for two-dimensional position sensing, *Res. Rep.*, HPL-92-19, HP Laboratories, Bristol, 1992. Available at <http://www.hpl.hp.com/techreports/92/HPL-92-19.pdf> (accessed Feb. 25, 2019).
6. **J. Dénes** and **A. D. Keedwell**, A new construction of two-dimensional arrays with the window property, *IEEE Trans. Inf. Theory*, **36**, No. 4, 873–876, 1990.
7. **T. Etzion**, Sequence folding, lattice tiling, and multidimensional coding, 2009 (Cornell Univ. Libr. e-Print Archive, arXiv:0911.1745).
8. **V. Horan** and **B. Stevens**, Locating patterns in the de Bruijn torus, *Discrete Math.*, **339**, No. 4, 1274–1282, 2016.
9. **G. Hurlbert** and **G. Isaak**, On the de Bruijn torus problem, *J. Comb. Theory, Ser. A*, **64**, No. 1, 50–62, 1993.
10. **S. A. Lloyd** and **J. Burns**, Finding the position of a subarray in a pseudo-random array, in *Cryptography and Coding III*, Clarendon Press, Oxford, 1993.
11. **C. J. Mitchell** and **K. G. Paterson**, Decoding perfect maps, *Des. Codes Cryptogr.*, **4**, No. 1, 11–30, 1994.
12. **C. J. Mitchell**, Aperiodic and semi-periodic perfect maps, *IEEE Trans. Inf. Theory*, **41**, No. 1, 88–95, 1995.
13. **K. G. Paterson**, New classes of perfect maps I, *J. Comb. Theory, Ser. A*, **73**, No. 2, 302–334, 1996.
14. **K. G. Paterson**, New classes of perfect maps II, *J. Comb. Theory, Ser. A*, **73**, No. 2, 335–345, 1996.
15. **W. C. Shiu**, Decoding de Bruijn arrays constructed by FFMS method, *Ars Comb.*, **47**, 33–48, 1997.
16. **J. Tuliani**, De Bruijn sequences with efficient decoding algorithms, *Discrete Math.*, **226**, No. 1–3, 313–336, 2001.

Dmitry A. Makarov
Alexey D. Yashunsky

Received October 30, 2018
Revised February 14, 2019
Accepted February 27, 2019