M. A. Kronrod, An optimal ordering algorithm without a field of operation,
Dokl. Akad. Nauk SSSR, 1969, Volume 186, Number 6, 1256–1258

Доклады Академии наук СССР
1969. Том 186, № 6

УДК 518.5

МАТЕМАТИКА

М. А. КРОНРОД

ОПТИМАЛЬНЫЙ АЛГОРИТМ УПОРЯДОЧЕНИЯ БЕЗ РАБОЧЕГО ПОЛЯ

(Представлено академиком И. Г. Петровским 25 XII 1968)

Легко показать, что нельзя упорядочить массив из n чисел меньше, чем за $Cn \lg n$ действий. Шелл построил алгоритм упорядочения за $Cn \lg n$ действий, но этот алгоритм использует рабочее поле длины порядка n. (Точнее: при наличии свободного поля длины k требуется $C(n^2/k + n \lg k)$ действий.) М. Я. Вайнштейн построил алгоритм упорядочения за $Cn \lg^2 n$ действий без рабочего поля.

В настоящей работе показано, что упорядочение без рабочего поля можно произвести за $Cn \lg n$ действий *.

Обозначения. Под словами массив ячеек подразумевается массив последовательно расположенных ячеек. Массивы ячеек обозначаются M, L, W и т. д.; M_i означает i-ю ячейку массива M, а SM_i — содержащееся в ней число. Участок массива M от M_i до M_j включительно обозначается $M[i:j]$.

Если в массиве M выполняется неравенство $SM_1 \geqslant SM_2 \geqslant \ldots \geqslant SM_n$, то мы говорим, что содержимое M упорядочено.

Если $SM_1 \geqslant SM_2 \geqslant \ldots \geqslant SM_k$ и $SM_{k+1} \geqslant SM_{k+2} \geqslant \ldots \geqslant SM_n$, то мы говорим, что содержимое M упорядочено для массива с границей k.

Всюду в дальнейшем используются логарифмы только по основанию 2, и потому для краткости мы будем писать $\lg n$ вместо $\log_2 n$.

В каждом пункте (в каждой лемме или теореме) имеется в виду существование программы, использующей ограниченное количество рабочих ячеек и выполняющей формулировку пункта. Множитель C, встречающийся в формулировке каждого пункта, является своим для каждого пункта.

Лемма 1. *Пусть имеются два массива M и L из n ячеек каждый, $SM_i = a_i$, $SL_i = b_i$ $(i = 1, 2, \ldots, n)$. Тогда содержимое этих массивов можно за Cn действий поменять местами, т. е. сделать так, чтобы было $SM_i = b_i$, $SL_i = a_i$.*

Лемма 2. *Содержимое массива M из n ячеек можно упорядочить за Cn^2 действий.*

Лемма 3. *Пусть имеется массив M из n ячеек и массив W, называемый вспомогательным, из k или больше ячеек, и пусть содержимое массива M упорядочено в два массива с границей k или $(n - k)$. Тогда можно за Cn действий упорядочить содержимое массива M, изменив при этом, вообще говоря, порядок расположения чисел внутри массива W.*

Доказательство. Пусть, например, граница равна k. Сначала по лемме 1 поменяем местами содержимое массивов $M[1:k]$ и $W[1:k]$. Кандидатами на первое место в массиве M являются числа, лежащие в ячейках M_{k+1} и W_1; то из них, которое больше, поменяем местами с содержимым M_1. Кандидатами на второе место являются: отвергнутый канди-

* Формально требуется одна рабочая ячейка для того, чтобы иметь возможность обменять два числа местами; в ЭВМ, у которой имеется, например, команда «сверка», этот обмен и, как следствие, все упорядочение производится вовсе без рабочих ячеек.

дат на первое место и число, лежавшее за бо́льшим из кандидатов на первое место. То из этих чисел, которое больше, меняем местами с содержимым M_2 и т. д. Когда произойдет обмен содержимого некоторого M_i с содержимым W_k, упорядочение будет закончено.

Случай, когда граница равна $(n-k)$, рассматривается аналогично, только начинать нужно не с кандидатов на первое, а с кандидатов на последнее место.

Лемма 4. *Содержимое массива M из n ячеек можно упорядочить за $C(n+k^2)$ действий, если содержимое массива $M[k+1:n]$ было упорядочено.*

Доказательство. По лемме 2 упорядочим содержимое массива $M[1:2k]$. Заметим, что старшие k чисел лежат теперь в массиве $M[1:k]$. Используя этот массив в качестве вспомогательного, упорядочим по лемме 3 содержимое массива $M[k+1:n]$. Теперь осталось только упорядочить по лемме 2 содержимое массива $M[1:k]$.

Теорема 1. *Пусть дан массив M из n ячеек, и пусть содержимое массива M упорядочено в два массива с границей l. Тогда можно за Cn действий упорядочить содержимое M.*

Доказательство. Положим $k=[\sqrt{n}]$; $p\equiv l(\operatorname{mod} k)$; $q\equiv(n-l)(\operatorname{mod} k)$. Назовем **зонами** массивы из k ячеек каждый: $Z^{(1)}=M[p+k+1:p+2k]$ — первая зона, $Z^{(2)}=M[p+2k+1:p+3k]$ — вторая зона, и т. д. Таким образом, мы разбили массив M на t зон, массив $M[n-q+1:n]$ и массив $M[1:p+k]$, который мы будем использовать как вспомогательный. Содержимое каждой зоны упорядочено.

Первое число каждой зоны назовем **признаком** зоны. Найдем зону с максимальным признаком и поменяем ее содержимое с содержимым первой зоны; потом найдем зону со вторым по величине признаком и поменяем ее содержимое с содержимым второй зоны и т. д.

Замечание. Теперь для каждой зоны $Z^{(i)}$ имеется не больше k чисел, меньших, чем признак $Z^{(i)}$, и лежащих в зонах с номерами, меньшими, чем i.

По лемме 3 упорядочим содержимое массива $M[p+k+1:p+3k]$, используя в качестве вспомогательного массива $M[1:p+k]$; потом по той же лемме упорядочим содержимое $M[p+2k+1:p+4k]$. Из замечания следует, что теперь содержимое $M[p+k+1:p+4k]$ будет упорядочено. Затем упорядочим содержимое $M[p+3k+1:p+5k]$, и т. д.

Так мы упорядочим содержимое массива $M[p+k+1:n-q]$. Так как содержимое массива $M[n-q+1:n]$ упорядочено, мы можем теперь по лемме 3 упорядочить содержимое массива $M[p+k+1:n]$, а затем, по лемме 4,— весь массив M.

Теорема 2 (основная). *Содержимое массива M из n ячеек можно упорядочить за Cn действий.*

Доказательство. При доказательстве мы будем использовать теорему 1. Сначала упорядочим содержимое массивов $M[1:2]$, $M[3:4]$ и т. д. (Все эти массивы, кроме, может быть, последнего, имеют длину, равную 2, а последний — меньше, либо равную 2.) Потом упорядочим содержимое массивов $M[1:4]$, $M[5:8]$ и т. д., потом — содержимое массивов по 8 ячеек, по 16 ячеек и т. д.

Теорема 3 (обобщение теоремы 2). *Пусть дан массив M из n ячеек, и пусть содержимое каждого из массивов $M[1:i_1]$, $M[i_1+1:i_2],\ldots$ $\ldots, M[i_{k-1}+1:n]$ упорядочено. Тогда можно за $Cn\lg k$ действий упорядочить содержимое массива M.*

Доказательство теоремы 3 аналогично доказательству теоремы 2. По теореме 1 упорядочивается сначала содержимое массивов $M[1:i_2]$, $M[i_2+1:i_4]$, $M[i_4+1:i_6]$ и т. д., потом — содержимое массивов $M[1:i_4]$, $M[i_4+1:i_8]$, $M[i_8+1:i_{12}]$ и т. д.

Для алгоритмов, существование которых доказывается теоремами 1—3, получаются большие множители C при реализации алгоритма на ЭВМ. Существуют другие алгоритмы с меньшими множителями. Основой для них всех является лемма 3. С ее помощью можно доказать следующую лемму (вообще говоря, эта лемма является частным случаем теоремы 2, но приводимый в ее доказательстве алгоритм имеет меньший множитель C).

Лемма 5. *Пусть имеется массив M из n ячеек и массив W, называемый вспомогательным, из n или больше ячеек. Тогда можно за $Cn \lg n$ действий упорядочить содержимое массива M, изменив при этом, вообще говоря, порядок расположения чисел в массиве W.*

Доказательство теоремы 2 теперь выглядит, например, так. Положим $k_1 = \left[\frac{n+1}{2}\right]$. По лемме 5 упорядочим содержимое массива $M[k_1+1:n]$, используя $M[1:k_1]$ в качестве вспомогательного. Положим $k_2 = \left[\frac{k_1+1}{2}\right]$. Используя $M[1:k_2]$ в качестве вспомогательного, упорядочим сначала содержимое $M[k_2+1:k_1]$ по лемме 5, а затем — содержимое $M[k_2+1:n]$ по лемме 3. Положим $k_3 = \left[\frac{k_2+1}{2}\right]$ и т. д. Так как упорядочим содержимое всего массива M за $Cn \lg n$ действий. Множитель C здесь значительно меньше, чем в первом доказательстве. Но и его можно уменьшить в 2 раза, если сделать только $p = \lg (\lg n)$ шагов, а затем, например, применить теорему 2 (с первым доказательством) к массиву $M[1:k_p]$ и упорядочить затем содержимое всего массива M по теореме 1.

Московский государственный университет
им. М. В. Ломоносова

Поступило
14 XII 1968