Секция 5

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

УДК 004.021

РЕАЛИЗАЦИЯ МАССОВЫХ ОПЕРАЦИЙ НА МНОГОМЕРНЫХ МАССИВАХ ДАННЫХ

А. Г. Банных

Для эффективной работы с информацией разработаны разнообразные структуры данных. Одна из самых распространённых из них — дерево. Широко известны квадродеревья [1], кд-деревья [2], обобщения дерева отрезков [3] и дерева Фенвика [4]. Существующие структуры данных для работы с многомерными массивами данных позволяют эффективно получать статистическую информацию и изменять *отдельные* элементы.

Цель настоящей работы заключается в том, чтобы исследовать возможность выполнения массовых обновлений — единообразного изменения *областей* данных. Ни одна из вышеперечисленных структур данных в чистом виде не предоставляет подобную функциональность. Подробнее с возникающими проблемами можно ознакомиться в [5].

Рассмотрим d-мерное пространство $\mathbb{Z}^d$. Пусть $\mathbf{x},\mathbf{y}\in\mathbb{Z}^d$; будем говорить, что $\mathbf{x}\leqslant\mathbf{y}$, если $x_k\leqslant y_k$ для всех $k\in\{1,2,\ldots,d\}$. Для любого $c\in\mathbb{Z}$ будем обозначать $\mathbf{c}=(c,c,\ldots,c)\in\mathbb{Z}^d$; для любых $\mathbf{x},\mathbf{y}\in\mathbb{Z}^d$, $\mathbf{x}\leqslant\mathbf{y}$, положим $P_{\mathbf{x},\mathbf{y}}=[x_1..y_1]\times[x_2..y_2]\times\ldots\times[x_d..y_d]$.

Рассмотрим абелеву группу $\langle G,+\rangle$ и отображение $A:P_{\mathbf{1},\mathbf{n}}\to G$, $n\in\mathbb{Z}$. Можно трактовать A как многомерный массив с множеством индексов $P_{\mathbf{1},\mathbf{n}}$ и множеством допустимых значений ячеек G.

Массовый запрос определим следующим образом: $\operatorname{get}(A,P)=\sum\limits_{\mathbf{x}\in P}A(\mathbf{x})$.

Массовое обновление: $\operatorname{update}(A,P,v)=A':P_{\mathbf{1},\mathbf{n}}\to G$, где

$$A'(\mathbf{x})=\begin{cases}A(\mathbf{x})+v, & \mathbf{x}\in P,\\ A(\mathbf{x}), & \mathbf{x}\notin P.\end{cases}$$

Определим оператор D_i следующим образом:

$$(D_iA)(\mathbf{x})=A(x_1,\ldots,x_i,\ldots,x_d)-A(x_1,\ldots,x_i-1,\ldots,x_d),$$

полагая $A(\mathbf{x})=0$ при $\mathbf{x}\notin P_{\mathbf{1},\mathbf{n}}$. Заменим элементы исходного массива $A(\mathbf{x})$ на линейные многочлены от d переменных $C(\mathbf{x})(\mathbf{z})=(D_1D_2\cdots D_dA)(\mathbf{x})\prod\limits_{i=1}^{d}(z_i-x_i)$.

Утверждение 1. Для области вида $P_{\mathbf{1},\mathbf{z}}$ массовый запрос сводится к массовому запросу на массиве C и подстановке переменных:

$$\operatorname{get}(A,P_{\mathbf{1},\mathbf{z}})=\operatorname{get}(C,P_{\mathbf{1},\mathbf{z}})(\mathbf{z}).$$

Утверждение 2. Для области вида $P_{\mathbf{x},\mathbf{n}}$ массовое обновление равносильно единичному обновлению в новом массиве:

$$\mathrm{update}(A, P_{\mathbf{x},\mathbf{n}}, v) \Leftrightarrow \mathrm{update}(C, P_{\mathbf{x},\mathbf{x}}, v \prod_{i=1}^{d} (z_i - x_i)).$$

Ослабленным назовем массовый запрос или обновление, для которого применимо утверждение 1 или 2 соответственно.

Утверждение 3. Произвольные массовые обновления и запросы выражаются не более чем через 2^d ослабленных.

Утверждения 1–3 позволяют свести задачу с массовыми обновлениями к хорошо изученной задаче с единичными обновлениями. Последнюю можно решать с использованием любых существующих структур данных. Это свойство позволяет делать выбор оптимальной структуры данных для каждой задачи отдельно.

При использовании популярных структур данных асимптотическая оценка на выполнение массовых операций составляет $\mathrm{O}(\log^d n)$. При малой размерности пространства эта оценка позволяет говорить об эффективности предлагаемого метода сведения задачи с массовыми обновлениями к задаче с единичными обновлениями. Свобода выбора структуры данных свидетельствует о гибкости метода.

Отметим, что не все задачи выполнения массовых операций на многомерных массивах данных можно сформулировать в терминах, использованных в данной работе. Вопрос об обобщении подхода на более широкие классы задач остаётся открытым.

ЛИТЕРАТУРА

1. *Романовский И. В.* Дискретный анализ. СПб.: Невский Диалект; БХВ-Петербург, 2008. 336 с.
2. *Bentley J. L.* Multidimensional binary search trees used for associative searching // Commun. ACM. 1975. V. 18. No. 9. P. 509–517.
3. *Препарата Ф., Шеймос М.* Вычислительная геометрия. Введение. М.: Мир, 1989. 478 с.
4. *Fenwick P. M.* A New Data Structure for Cumulative Frequency Tables // Software: Practice and Experience. 1994. V. 24. No. 3. P. 327–336.
5. `http://is.ifmo.ru/papers/2011-bachelor-bannykh/` — Банных А. Г. Применение деревьев для реализации массовых операций на многомерных массивах данных [Электронный ресурс]. 2011.

УДК 518.517

АНАЛИЗ АЛГОРИТМОВ ИСПОЛНЯЕМЫХ ФАЙЛОВ ВНУТРИ ПРОГРАММНОЙ ЭВМ С ИЗМЕНЯЕМОЙ СПЕЦИФИКАЦИЕЙ

А. С. Бурлаков

Одним из этапов анализа программного обеспечения является низкоуровневый анализ алгоритмов, скомпилированных в исполняемый файл. Несмотря на то, что данная проблема является довольно популярной, для неё не существует тривиального решения, а имеется набор инструментов и методов [1, с. 3]. Одним из способов решения данной проблемы является использование дизассемблера, чего, однако, часто бывает недостаточно в силу того, что пользователя интересует, каким образом программа взаимодействует с внешними устройствами [2, с. 30–38].