

Теорія алгоритмів

Практичне завдання №6

“Піраміди”

Завдання

В даній роботі необхідно розв'язати наступну задачу визначення послідовності медіан для заданого вхідного масиву. Нагадаємо, що медіаною для масиву називається елемент, який займає середнє положення у відсортованому масиві. Так, якщо кількість елементів у масиві непарна, то медіана одна та індекс її у відсортованому масиві визначається як $\lfloor n/2 \rfloor$ (де n — розмір вхідного масиву). Якщо кількість елементів у масиві парна, то медіан буде дві та їх індекси визначаються за формулами $\lfloor n/2 \rfloor$ та $\lfloor n/2 \rfloor + 1$.

Задача формулюється наступним чином. Нехай заданий вхідний масив $A = [x_1, \dots, x_N]$. Припустимо, що елементи масиву поступають на вхід програми послідовно: в кожний момент часу розглядається новий елемент x_i . Необхідно для кожного i (від 1 до N) визначити медіану підмасиву $A' = [x_1, \dots, x_i]$, тобто медіану для масиву елементів, які були отримані програмою на даний момент часу. Необхідно розв'язати цю задачу, використовуючи структури даних пірамід і так, щоб кожна медіана визначалась за час $O(\log(i))$.

Цю задачу можна розв'язати, використовуючи дві піраміди (heap) наступним чином.

- Позначимо через H_{low} незростаючу піраміду (max-heap), яка буде містити елементи меншої половини масиву (тобто такі елементи, які у відсортованому поточному елементі A' будуть розташовуватись у першій, меншій половині масиву).
- Позначимо через H_{high} неспадну піраміду (min-heap), яка буде містити елементи більшої половини масиву (тобто такі елементи, які у відсортованому поточному елементі A' будуть розташовуватись у другій, більшій половині масиву).

Тепер розглянемо роботу процедури, яка розв'язує поставлену задачу із використанням двох наведених пірамід. Нехай додається черговий елемент x_i . На поточний момент сумарна кількість елементів, які зберігаються в обох пірамідах, становить $(i-1)$. Наступні кроки, які ми повинні виконати:

1. Визначимо, в яку піраміду (H_{low} або H_{high}) потрібно додати новий елемент. Якщо x_i менше ніж найбільший елемент з H_{low} (тобто новий елемент буде розташовуватись в меншій поточній половині), то додаємо його у цю піраміду. В іншому випадку додаємо елемент в піраміду H_{high} .
2. В кожний момент часу, тобто на кожній ітерації роботи алгоритму, повинен зберігатись наступний інваріант: кількість елементів в піраміді H_{low} не повинна відрізнятись від кількості елементів в H_{high} не більше ніж на одиницю. Під час виконання попереднього етапу цей інваріант може порушитись. Тому тепер необхідно відновити даний інваріант: якщо у піраміді H_{low} елементів більше на 2 за H_{high} , то визначаємо найбільший елемент з H_{low} і вставляємо його у H_{high} ; якщо кількість елементів у H_{high} більше на 2 за H_{low} , то визначаємо найменший елемент з H_{high} і вставляємо його у H_{low} . Зрозуміло, що після кожної вставки нового елементу в піраміду необхідно перевіряти властивість піраміди: для H_{low} - властивість незростаючої піраміди, для H_{high} - властивість неспадної піраміди.
3. Визначити медіану для поточного масиву $A' = [x_1, \dots, x_i]$:
 - Якщо кількість елементів у A' парна, то після збереження інваріанту у пункті 2, кількість елементів у пірамідах H_{low} та H_{high} буде рівною. Тому одна медіана буде найбільшим елементом H_{low} , а інша медіана — найменшим елементом H_{high} .
 - Якщо кількість елементів у A' непарна, то єдина медіана буде знаходитись у тій піраміді, в якій кількість елементів буде більше (на одиницю) за кількість в іншій. Тому, якщо кількість елементів у H_{low} більше за H_{high} , то медіана — це найбільший елемент з H_{low} . Інакше медіана — найменший елемент з H_{high} .

Наведений алгоритм використовує процедури `extract_max` незростаючої піраміди H_{low} та

extract_min неспадної піраміди H_{high} , які виконуються за час $O(\log N)$, де N — розмір піраміди. Тому на кожній ітерації № i для поточного масиву $A' = [x_1, \dots, x_i]$ час роботи наведеної процедури становитиме $O(\log(i))$.

Формат вхідних/вихідних даних

Розроблена програма повинна зчитувати вхідні дані з файлу заданого формату та записувати дані у файл заданого формату. Вхідний файл представляє собою текстовий файл із $N+1$ рядків, де N — це розмірність вхідного масиву A . Першим записом є число — кількість елементів в масиві; наступні N записів містять елементи вхідного масиву.

Вихідний файл представляє також текстовий файл із N рядків, де кожен рядок i містить медіани для вхідного підмасиву $[x_1, \dots, x_i]$. Якщо медіана одна, то в рядку буде одне число; якщо медіани дві, то вони записуються через пробіл.

До документу завдання також додаються приклади вхідних і вихідних файлів різної розмірності.

Нижче наведені приклади вхідного та вихідного файлу для $N = 10$.

Вхідний файл	Вихідний файл
10	6
6	6 10
10	7
7	6 7
1	6
4	6 7
8	6
3	6 7
9	6
5	5 6
2	