

SPb HSE, ПАДИИ, 1 курс, осень 2024/25

Практика по алгоритмам #5

Stable-sorts, хеш-таблицы

3 октября

Собрано 30 сентября 2024 г. в 14:07

---

## Содержание

1. Stable-sorts, хеш-таблицы	1
2. Домашнее задание	2
2.1. Дополнительная часть . . . . .	2

# Stable-sorts, хеш-таблицы

## 1. Тройки чисел

Как стабильно отсортировать тройки чисел?

## 2. Ревизия сортировок

Какие сортировки вы уже знаете? Какие из них стабильны, какие нет?

## 3. За сколько мы умеем сортировать числа?

$n = 10^5$ , целые числа от 1 до  $10^9$ .

## 4. Интерфейс хеш-таблицы

Что и за сколько умеет `set` реализованный на хеш-таблице?

Что и за сколько умеет `map` реализованный на хеш-таблице?

## 5. Память в хеш-таблицах для `set`

Мы хотим свой `set`. Посчитайте, сколько байт памяти нужно хеш-таблице на списках для хранения  $n$  интов в  $n$  списках? Посчитайте, сколько памяти нужно хеш-таблице с открытой адресацией для хранения  $n$  интов в массиве размера  $1.5n$ ? Кто лучше?

## 6. Что будет с хеш-таблицей?

Мы хотим хранить  $n$  интов в хеш-таблице размера  $N$ . Что будет со списками? с открытой адресацией? (a)  $N = n$ ; (b)  $N = \frac{n}{2}$ ; (c)  $N = 1.5n$ .

## 7. Оценка хеш-функций

Покажите, что хеш функция  $x \rightarrow x \bmod N$ , где  $N$  случайное простое в диапазоне  $[n, 2n)$

(a) хороша для хеш-таблицы на списках;

(b) плохо работает для хеш-таблицы с открытой адресацией с шагом  $+1$ .

## 8. Патчим списки

Различные числа стоят по кругу. Нужно за  $\mathcal{O}(1)$  обрабатывать запросы: стереть число  $x$ ; добавить в круг число  $x$  после числа  $y$ . Гарантируются, что все числа различны.  $x \leq 10^{18}$ .

## 9. BucketSort

Пусть есть случайные равномерно распределённые целочисленные данные.

Придумайте аналог CountSort-а, который использует  $\Theta(n)$  доп памяти.

## 10. Суффиксный массив

Дана строка длины  $n$ , отсортируйте все её суффиксы за  $\mathcal{O}(n^2)$ . Научитесь с помощью суффиксного массива от текста искать вхождение строки  $s$  в текст.

## 11. d-куча

За сколько будет работать куча, которая ветвится на  $d$  направлений?

## 12. (\*) Быстрый сортировок не бывает?

Докажите, что  $n$  объектов, для которых определено только сравнение, нельзя сортировать быстрее чем  $n \log n$ ?

## Домашнее задание

### 1. (2) Амортизированное время

С изначально пустой хеш-таблицей собираются сделать какие-то  $n$  операций `add/del/find`. Опишите модификацию хеш-таблицы, которая умеет `add` и `del` за  $\mathcal{O}(1)$  на каждую операцию в худшем случае, и матожидание суммарного времени работы всех  $n$  операций есть  $\mathcal{O}(n)$ .

*Подсказка: за сколько работают уже известные вам хеш-таблицы?  
Помните ли вы что-либо про ленивое удаление?*

### 2.1. Дополнительная часть

#### 1. (2) Сортировка строк

Мы уже умеем `radix-sort`-ом сортировать  $n$  строк длины  $k$  над алфавитом размера  $\leq n$  за  $\mathcal{O}(nk)$ . Научитесь теперь сортировать  $n$  строк разной длины.

Пусть суммарная длина строк  $L$ , отсортируйте строки за  $\mathcal{O}(L)$ .

#### 2. (2) Различные элементы

*Задача:* дан массив, проверить, что все элементы различны.

Пусть для работы с элементами массива нам доступна только операция сравнения. Доказать оценку  $\Omega(n \log n)$  на число сравнений.