МФТИ

Алгоритмы и структуры данных, осень 2022

Домашнее задание №04. Хеш-таблицы

1. (1 балл) Дан массив чисел длины n, а также число k. В каждом подотрезке массива длины k найдите количество различных чисел, то есть количество чисел, встречающихся на отрезке хотя бы один раз. Асимптотика: $O(n)$ в среднем.

2. (1 балл) Дан массив чисел длины n и число C. Определите, можно ли выбрать три элемента массива так, чтобы их сумма равнялась C. Асимптотика: $O(n^2)$ в среднем.

3. (1 балл) Назовём два массива натуральных чисел $a_1, \ldots, a_n$ и $b_1, \ldots, b_n$ изоморфными, если существует отображение $\varphi : \mathbb{N} \to \mathbb{N}$, такое что $\varphi(a_i) = b_i$ для всех i, а также существует $\psi : \mathbb{N} \to \mathbb{N}$, такое что $\psi(b_i) = a_i$ для всех i. По данным массивам определите, изоморфны ли они. Асимптотика: $O(n)$ в среднем.

4. (3 балла) Научимся хешировать кортежи чисел длины d. Кортежу $(a_0, \ldots, a_{d-1})$ сопоставим число $h_b(a) = (a_0 + a_1 b + \ldots + a_{d-1} b^{d-1}) \bmod p$, где p — фиксированное простое число, а b — случайное. Покажите, что сравнение $h_b(a) = v$ (как уравнение относительно b) имеет не более d решений в поле $\mathbb{Z}_p$. Пусть теперь $H_b(a) = h_b(a) \bmod m$, а $\mathcal{H} = \{H_b(\cdot) \mid b \in [1, p-1]\}$. Покажите, что $\mathbb{P}(h(x) = h(y)) \leqslant \frac{d}{m}$ для фиксированной пары различных кортежей x и y. Выведите отсюда, что при таком способе хеширования среднее время работы всех операций вырастет до $O^*(d \cdot \alpha)$, где α — показатель загруженности хеш-таблицы.