МФТИ, ФПМИ

Алгоритмы и структуры данных, осень 2021

Семинар №02. Бор, алгоритм Ахо—Корасик

1. По префикс-функции строки постройте её z-функцию.
2. Дана строка s. С помощью линейного предподсчёта научитесь отвечать на запросы вида “является ли подстрока $s_l \dots s_r$ палиндромом?” за $O(1)$.
3. В множестве S лежат n чисел: $x_1, \dots, x_n$. Для каждого из чисел $\{y_i\}_{i=1}^q$ найдите $\max\limits_j (y_i \oplus x_j)$. Проделайте то же для минимума. В предположении, что все числа — целые и лежат в отрезке $[0, 2^k - 1]$, отвечайте на каждый запрос за $O(k)$.
4. К множеству S поступают запросы двух видов: добавить строку в S; сообщить k-ю строку S в лексикографическом порядке (k — параметр запроса). Отвечайте на запрос за линейное время от длины строки. Как с помощью аналогичной техники отсортировать заданный список строк?
5. Как реализовать алгоритм Ахо—Корасик на динамически расширяющемся множестве строк? Используйте идею разложения n по степеням двойки.
6. Дана строка s с не более чем k знаками вопроса. Вхождением s в текст t назовём подстроку t, которая совпадает с s во всех символах, кроме вопросов. Найдите все вхождения s в t за время $O((|s| + |t|) \cdot k)$.
7. Рассмотрим алфавит из 4 букв и n строк в нём: $s_1, \dots, s_n$, причём $|s_i| \leqslant m$ для всех i. Слово t назовём хорошим, если в нём можно выделить несколько подстрок, каждая из которых равна какому-нибудь s_i, и все символы при этом находятся хотя бы в одной из выделенных подстрок (то есть t покрыта словарным словами). Найдите число хороших строк длины k.
8. Задан фиксированный словарь, каждое слово имеет свой вес. Поступают запросы двух видов: “изменить вес i-го словарного слова” и “найти словарное слово с максимальным весом, которое входит в текст t как подстрока”. Отвечайте на запросы первого типа за $O(\log^2 n)$, где n — суммарная длина слов словаря, а на запросы второго типа — за $O(|t| \log^2 n)$.