

13 Строки: алгоритм КМП, Ахо-Корасик

1. Словари offline

Даны словарь (конечное множество слов) и текст.

- Для каждого слова из словаря определить, входит ли оно в текст.
- Для каждого слова из словаря найти число вхождений в текст.

2. Словари online

Даны словарь (конечное множество слов) и текст. Обновлять ответ online при добавлении символа в конец текста.

- Пересчитать суммарное число вхождений слов из словаря в текст за $\mathcal{O}(1)$.
- Пересчитать множество всех вхождений слов из словаря в текст за $\mathcal{O}(1 + |\Delta A|)$, где ΔA — приращение ответа после добавления очередного символа. Вхождение слова — это номер слова и позиция в тексте, с которого его можно считать.

3. (!) Ахо-Корасик и замыкание до автомата

На лекции мы научились считать суффиксные ссылки и находить вхождения строк в текст. Для этого мы в каждый момент поддерживали позицию в боре, и в случае когда перехода по нужному символу не было, происходит откат по суффиксным ссылкам.

Однако, можно замкнуть бор до автомата, и в таком случае переход по букве c всегда будет выглядеть как $v \rightarrow v.go[c]$, без откатываний. Опишите как построить замыкание до автомата.

4. (!) Избегаемость шаблонов («Вирусы»)

Дан словарь слов суммарной длины L над алфавитом Σ . За время $\mathcal{O}(L \cdot |\Sigma|)$ определите, существует ли бесконечная строка, не содержащая ни одно словарное слово как подстроку.

5. XOR \rightarrow max

Дан массив a длины n . Найдите пару $a_i, a_j: a_i \wedge a_j = \max$. Время $\mathcal{O}(n \log M)$, где $M = \max(a_1, \dots, a_n)$.

6. XOR $\geq k$

Дан массив a длины n и число k . Длина чисел в массиве $\mathcal{O}(1)$. За время $\mathcal{O}(n)$ посчитайте количество

- пар индексов таких, что побитовый XOR элементов по этим индексам $\geq k$.
- отрезков последовательности, побитовый XOR всех чисел из которых $\geq k$.

7. Количество строк

Дан словарь и число n . Посчитайте количество строк длины n над алфавитом $\{a, b\}$, которые не содержат ни одного словарного слова, как подстроку. Так как ответ может быть очень большим, выведите количество по модулю $M = 10^9 + 9$.

- Время и память $\mathcal{O}(nL)$, где L — суммарная длина слов в словаре.
- (*) Время $\mathcal{O}(nL)$, память $\mathcal{O}(L)$.