

Алгоритмы. HW#12

Тураев Тимур, 504 (SE)

1 Найти рефрен

Построим по строке сжатое суффиксное дерево. Легко доказать, что в дереве будет линейное число внутренних вершин (листьев ровно $n + 1$, а каждая из внутренних вершин имеет по крайней мере 2 потомка, то есть прибавляет к своему поддереву по крайней мере 1 лист. Значит всего вершин – $O(n)$).

Запустим DFS, посчитаем число листьев в для каждой вершины – это число будет означать число вхождений строки, ассоциированной с вершиной.

Ну и все, задача решена – для каждой строки (понятно что на ребрах останавливаться нет смысла, всегда можно увеличить подстроку не изменив число вхождений пройдя до ближайшей вершины) мы знаем ее длину и число вхождений. Запускаем DFS, считаем максимум и выдает ответ.

Время, понятно, линейно.

3 Призрак Вася.

Для наглядности рассмотрим строку, которую может видеть Вася: *aaaaaxq*. Возможные исходные строки: *qxa~~aa~~* или *qx~~aa~~*.

Теперь ясно, что задачу можно переформулировать так: найти в заданной строке все префиксы-палиндромы четной длины. Это и будут концы исходной строки и отражение в зеркале. По этим данным легко восстановить все возможные исходные строки.

Как найти все такие префиксы? Легко проверить (и доказать тоже легко), что, если префикс длины k равен суффиксу той же длины у реверснutoй строки, то этот префикс является палиндромом.

Отсюда решение: разворачиваем строку, строим по ней суффиксное дерево, дальше идем по всем четным префиксам исходной строки и спускаемся по соответствующим символам в дереве. Если остановились в какой-то вершине, в которой заканчивается какой-то суффикс, то мы нашли тот самый префикс, равный суффиксу перевернутой строки – значит он палиндром. Продолжаем идти по ветке (и четным префиксам) дальше, пока можно, чтобы найти все возможные слова.