

Лабораторная работа № 5 по курсу дискретного анализа: Суффиксные деревья

Выполнил студент группы 08-310 МАИ *Филиппов Владимир*.

Условие

Кратко описывается задача:

1. Линеаризация циклической строки
2. Суффиксное дерево

Метод решения

Программа состоит из одного файла. Присутствует структура хранения узла дерева, который хранит индекс начала и конца подстроки текста, контейнер ссылок на детей и суффиксную ссылку. Построение дерева основано на алгоритме Укконена.

Описание программы

Разделения по файлам нет, все хранится в одном файле main.cpp. В наличии, структура для хранения узла дерева, сама структура дерева, содержащая реализацию алгоритма Укконена.

Дневник отладки

Столкнулся с несколькими багами при построении дерева. Неправильно закидывал суффиксные ссылки.

Производительность

Size (10^3)	Z Blocks (ms)
100	73
500	420
1000	1356
2500	2167
5000	6355

По рисунку 1 видим, что сложность - линейная. Оно и понятно. Алгоритм построения суффиксного дерева, а также проход по первым лексикографически узлам работает линейно от длины текста.

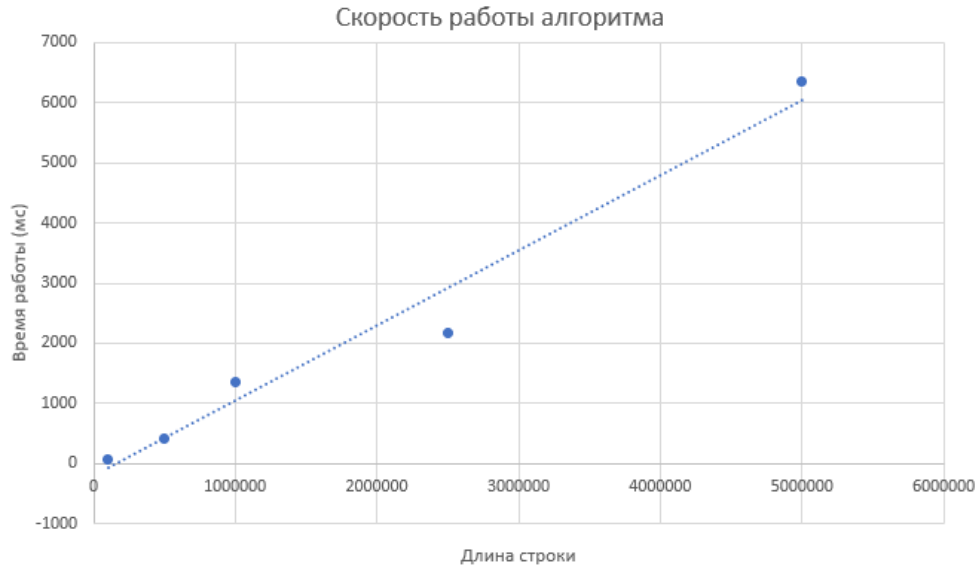


Рис. 1: Сравнение производительности

Забавно, что время работы алгоритма сильно зависит от входной строки. Так, в зависимости от данных время исполнения могло отличаться в два и более раз. Предполагаю, что это связано с особенностью построения суффиксного дерева: чем меньше у нас строка содержит повторений, тем меньше мы перемещаемся по суффиксным ссылкам, а следовательно дольше работает алгоритм.

Недочёты

Алгоритмически не решенных проблем нет. Можно лишь добавить архитектурного шарма: шаблонов, указателей и другого.

Выводы

Суффиксные деревья – очень полезная структура данных, так как при их помощи можно реализовать линейный поиск вхождений строк в текст с его предобработкой. Кроме этого на основе суффиксных деревьев основано множество алгоритмов на строках, таких как, например, реализовал я.

При выполнении данной л.р. я узнал про суффиксные деревья, написал реализацию алгоритма Укконена и решил задание из лабораторной работы. Надеюсь, что знание суффиксных деревьев поможет мне в дальнейшей карьере!