

Лабораторная работа № 5 по курсу дискретного анализа: суффиксные деревья

Выполнил студент группы 08-307 МАИ *Дегтярев Денис Андреевич*.

Условие

Найти образец в тексте используя статистику совпадений.

Формат ввода

На первой строке располагается образец, на второй — текст.

Формат вывода

Последовательность строк содержащих в себе номера позиций, начиная с которых встретился образец. Строки должны быть отсортированы в порядке возрастания номеров.

Метод решения

Находим образец подстроки в тексте с помощью статистики совпадений

Описание программы

Строим суффиксное дерево по паттерну. Затем вызываем метод GetMatchStatistics для того, чтобы построить статистику для вершин. Если данная статистика равна размеру паттерна, значит выводим индекс данной статистики(соответствует букве).

Дневник отладки

Все прошло с первого раза(удивлен)

Тест производительности

```
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ ./a.out < file.txt
Pattern size: 3
Text size: 7
Time: 6.1e-05
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ g++ main.cpp
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ ./a.out < file.txt
Pattern size: 6
Text size: 7
Time: 9.9e-05
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ g++ main.cpp
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ ./a.out < file.txt
Pattern size: 3
Text size: 67
Time: 0.000112
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ ./a.out < file.txt
Pattern size: 53
Text size: 67
Time: 0.000106
● denis@denis-Vivobook-ASUSLaptop-M6500QH-M6500QH:~/code/diskran/5$ ./a.out < file.txt
Pattern size: 53
Text size: 130
Time: 9e-05
```

Недочёты

Возможно, можно еще оптимизировать по скорости и убрать пару ненужного кода

Выводы

Выполнив лабораторную работу №5 по курсу «Дискретный анализ», я изучил различные методы построения суффиксных деревьев и алгоритмы, работающие с помощью суффиксных деревьев, и убедился, что они помогают в оптимизации решений различных задач, связанных с поиском подстроки в строке, благодаря представлению текста в более компактном виде.