Лабораторная работа № 5 по курсу дискретного анализа: Суффиксные деревья

Выполнил студент группы 08-207 МАИ Лебедев Иван.

Условие

1. Необходимо реализовать алгоритм Укконена построения суффиксного дерева за линейное время. Построив такое дерево для некоторых из входных строк, необходимо воспользоваться полученным суффиксным деревом для решения своего варианта задания.

Алфавит строк: строчные буквы латинского алфавита (т.е., от а до z).

2. Найти в заранее известном тексте поступающие на вход образцы с использованием суффиксного массива. Формат входных и выходных данных, а также примеры аналогичны указанным во первом задании.

Метод решения

К первой строке входных данных, прибавляется терминальный символ и строится суффиксное дерево по алгоритму Укконена. По полученному дереву строится суффиксный массив, в котором затем происходит поиск поступающих на вход слов в дереве.

Описание программы

1. Заголовочные файлы:

TSufTree.h(Объявление дерева и функций) TSufArray.h(Объявление суффиксного массива и функций)

2. Срр файлы:

TSufArray.cpp(Описание функций объявленных в TSufArray.h) TSufTree.cpp(Описание функций объявленных в TSufTree.h) main.cpp(Считывание данных, управление программой)

Дневник отладки

Программа не проходила тест из-за неверного вывода результата.

Выводы

Алгоритм Укконена - универсальный алгоритм для разных задач, связанных с поиском в строке. Он основан на алгоритме построения дерева с кубической сложностью, но с помощью различных улучшений работает за линейную сложность. Все эти модификации делают алгоритм Укконена сложным для понимания, однако очень эффективным в практике.