

## Лабораторная работа №4 по курсу Дискретного Анализа:

### Строковые алгоритмы

Выполнил студент группы М8О-207-21 МАИ *Кажекин Денис*.

#### Условие:

Необходимо реализовать один из стандартных алгоритмов поиска образцов для указанного алфавита.

**Вариант алгоритма:** Поиск одного образца основанный на построении Z-блоков.

**Вариант алфавита:** Слова не более 16 знаков латинского алфавита (регистронезависимые).

Запрещается реализовывать алгоритмы на алфавитах меньшей размерности, чем указано в задании.

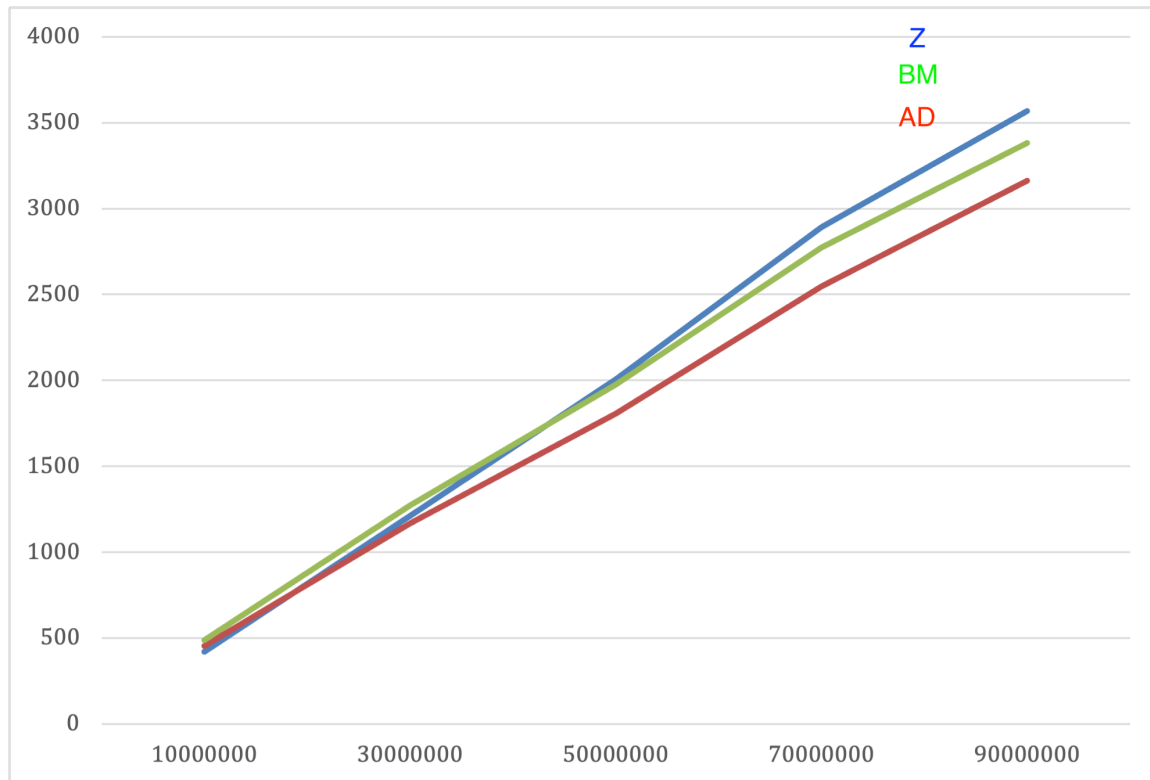
#### Метод решения/Описание алгоритма:

Для выполнения задачи нулевым шагом было необходимо реализовать функцию, которая будет вычислять Z-функцию для каждого символа входного текста за линейное время. Это достигается путем определения самого правого Z-блока, который ускоряет нахождение Z-функции. Также была реализована функция "trim", которая предобрабатывает строку, удаляя лишние пробелы между словами, а также символы переноса строки. Алгоритм состоял в следующем: нам необходимо считать все строки из входного файла, обработав их функцией "trim" и создать путем конкатенации итоговую строку для нахождения ответа вида: "P\$S", где P – паттерн, "\$" – разделяющий символ, S – строки входного файла. Также, чтобы учитывать паттерны, которые располагаются на нескольких строках, я считал количество слов на каждой считанной строке, и при проходе по сконкатенированной строке P\$S считал слова, и если значение Z – функции для символа было равно длине паттерна, то в зависимости от номера слова выводился определенный номер строки и слова. (для подсчета слов было необходимо реализовать массив префиксных сумм).

#### Дневник отладки

Была единственная ошибка, которая была обнаружена на тесте номер 7. Там выяснилось, что мой алгоритм неверно работал с пустыми строками, что сказывалось на неверной конкатенации строк (добавлялся ненужный пробел, что вело к неверному подсчету Z – функции). Я это исправил, просто добавив проверку на пустоту входной строки после чего программа успешно зашла в чекер.

## Тест производительности



Выше можно наблюдать сравнительную графику производительностей алгоритмов AD, BM и Z блоков. Z блоки выигрывают на меньших объемах данных, но проигрывают на больших. По вертикали скорость работы в миллисекундах, а по горизонтали объемы входных данных.

## Недочёты

Не обнаружены

## Выводы

Написав лабораторную работу, я изучил строковый алгоритм Z – блоков. Алгоритм, основанный на Z – блоках применим во всех задачах нахождения паттерна в заданном тексте. Сложностей при программировании не возникло.