

Лабораторная работа № 4 по курсу дискретного анализа: строковые алгоритмы

Выполнил студент группы 08-207 МАИ *Боев Савелий*.

Условие:

Необходимо реализовать поиск одного образца в тексте с использованием алгоритма Z-блоков. Алфавит — строчные латинские буквы.

Метод решения

Функция **Zfunction** реализует вычисление Z-функции для строки **s**. Z-функция определяет наибольшую длину префикса подстроки, начинающейся с позиции **i**, который совпадает с префиксом всей строки. Z-функция позволяет нам найти все вхождения образца в текст.

В функции **main** считывается строка **text**, в которой будет выполняться поиск, и строка **pattern**, которую мы хотим найти в тексте.

Затем, к строке **text** добавляется разделительный символ '\$', за которым следует строка **pattern**. Это сделано для того, чтобы поиск был возможен в любой позиции текста.

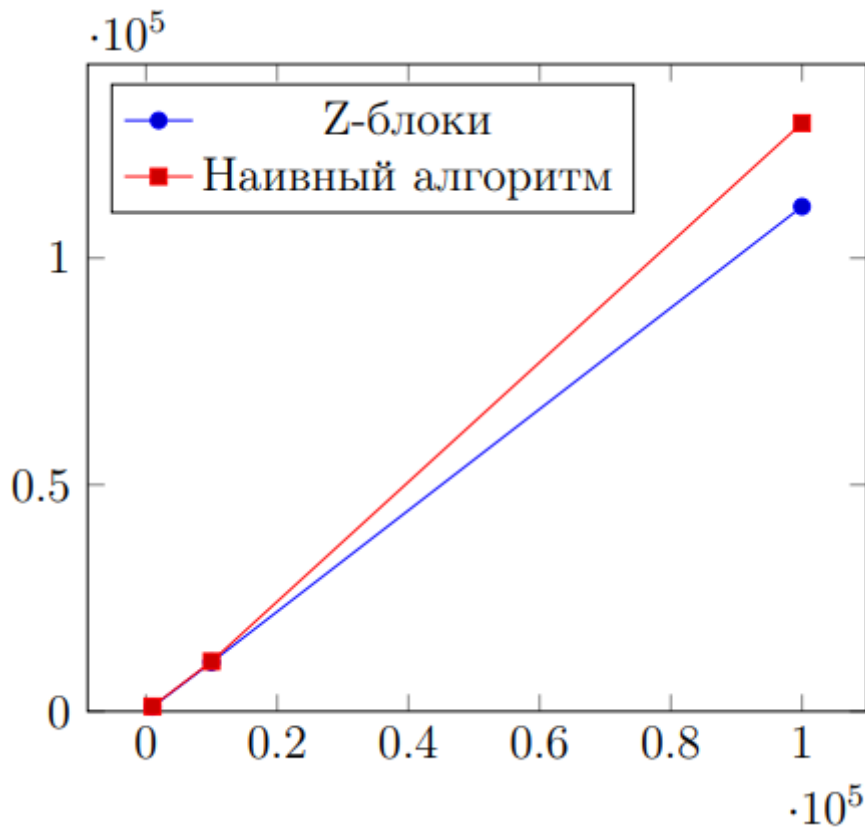
Вызывается функция **Zfunction**, которая вычисляет Z-функцию для модифицированной строки **text**.

Затем проходим по значениям Z-функции и проверяем, равно ли значение Z-функции в позиции **i** длине **pattern**. Если равно, то значит образец найден в тексте, и мы выводим позицию **i - pattern.size() - 1**, где **pattern.size()** - длина образца. Это позволяет получить начальную позицию образца в исходном тексте.

Описание программы

Программа состоит из одного файла, в котором реализован алгоритм поиска подстроки в строке с помощью z-блоков. Основные этапы работы программы: цикл на построение z-функции, цикл на поиск вхождения.

Тест производительности



Алгоритм Z-блоков оказался немного быстрее наивного алгоритма, так как длина образца во всех тестах была сравнительно небольшой, однако стоит понимать, что с увеличением длины образца будет увеличиваться и разница во времени работы алгоритмов.

Выводы

В данной лабораторной работе, я реализовал поиск всех вхождений образца в тексте с помощью алгоритма z-блоков. С помощью z-функции можно построить префикс-функцию, которая используется для алгоритма Кнута-Морриса-Пратта