

Лабораторная работа № 4 по курсу дискретного анализа: Поиск образца в строке

Выполнил студент группы 8О-207Б-20 МАИ *Чекменёв Вячеслав Алексеевич*.

Условие

1. **Упрощённый вариант:** поиск одного образца в тексте с помощью алгоритма Z-блоков.
2. **Алфавит:** строчные латинские буквы.

Метод решения

Объявим две функции - `zfunc()` и `findentries()`, в первой вычислим `z` функцию для строки вида `<pattern>@<text>`. Во второй пройдемся по строке и найдем индексы, в которых `z` функция равна длин паттерна.

Дневник отладки

1. Сначала написал программу с простой `z` функцией, которая работает за квадрат
2. Потом переписал, как на лекции

Тест производительности

Сравнение производительности производилось с наивным алгоритмом поиска подстроки. Время выполнения в миллисекундах.

Кол-во строк	Z-блоки	Наивный алгоритм
1000	856	1697
10000	8670	14975
100000	80367	199760

Исходя из результатов можно выяснить, что алгоритм работает значительно быстрее наивного алгоритма поиска, так как сложность нашего алгоритма $O(n+k)$

Выводы

После выполнения данной ЛР я познакомился с простейшим алгоритмом поиска образца в строке, который работает за линейное время.

Написание действительно оказалось простым, так как ввод был стандартный в отличие от вариантов не на оценку 3.

Плюсы данного алгоритма - сложность (асимптотическая) $= O(n+k)$, где k - длина образца, n - длина текста) на уровне с другими (более сложными) алгоритмами, легкость

написания.

Минусы - не так быстр как, например, алгоритм Бойера-Мура, так как не использует мелкие усечения по типу правил хорошего суффикса и плохого символа.