

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МО ЭВМ

ОТЧЕТ
по лабораторной работе №4
по дисциплине «Построение и анализ алгоритмов»
ТЕМА: АЛГОРИТМ КНУТА-МОРРИСА-ПРАТТА

Студент гр. 3388

Потоцкий С.С.

Преподаватель

Жангиров Т.Р.

Санкт-Петербург

2025

Цель работы

Изучить эффективный алгоритм Кнута-Морриса-Пратта, осуществляющий поиск подстроки в строке: его реализацию, асимптотику по времени работы, асимптотику по требуемой памяти. С помощью данного алгоритма решить ряд задач.

Задания.

Задание №1.

Реализуйте алгоритм КМП и с его помощью для заданных шаблона P ($|P| \leq 15000$) и текста T ($|T| \leq 5000000$) найдите все вхождения P в T .

Входные данные:

Первая строка - P

Вторая строка - T

Выходные данные:

Индексы начал вхождений P в T , разделенных запятой, если P не входит в T , то вывести -1

Пример входных данных:

ab

abab

Пример выходных данных:

0,2

Задание №2.

Заданы две строки A ($|A| \leq 5000000$) и B ($|B| \leq 5000000$).

Определить, является ли A циклическим сдвигом B (это значит, что A и B имеют одинаковую длину и A состоит из суффикса B , склеенного с префиксом B). Например, defabc является циклическим сдвигом abcdef.

Входные данные:

Первая строка - A

Вторая строка - B

Выходные данные:

Если A является циклическим сдвигом B, индекс начала строки B в A, иначе вывести -1. Если возможно несколько сдвигов вывести первый индекс.

Пример входных данных:

defabc

abcdef

Пример выходных данных:

3

Описание алгоритма

Первое задание

Реализует классический алгоритм Кнута-Морриса-Пратта (КМП) для поиска одного шаблона в тексте.

Ввод данных

Считывается строка-образец (pattern) и строка-текст (text).

Вычисление префикс-функции

Функция computeLPS строит префикс-функцию (массив lps), которая для каждой позиции шаблона определяет длину наибольшего собственного префикса, совпадающего с суффиксом.

Поиск всех вхождений шаблона

Функция kmpSearch проходит по тексту, используя префикс-функцию для эффективного поиска всех вхождений шаблона. При несовпадении символов алгоритм не возвращается к началу шаблона, а использует lps для смещения.

Вывод результата

В основной функции выводятся все позиции начала вхождений шаблона в текст (через запятую) либо -1, если вхождений нет.

Второе задание

Реализует алгоритм Кнута-Морриса-Пратта (КМП) для поиска первого вхождения шаблона и проверки циклического сдвига строк.

Ввод данных

Считываются две строки A и B.

Вычисление префикс-функции

Функция computeLPS строит префикс-функцию (массив lps) для шаблона.

Поиск первого вхождения шаблона

Функция kmpSearch ищет первое вхождение строки A в строке B+B с помощью КМП.

Проверка на циклический сдвиг

Функция isCyclicShift определяет, можно ли получить строку A циклическим сдвигом строки B. Для этого ищется A в B+B. Если найдено, вычисляется величина сдвига.

Вывод результата

В основной функции выводится величина сдвига, если строки являются циклическими сдвигами, иначе -1.

Оценка сложности алгоритма:

Время работы

Препроцессинг (построение префикс-функции): $O(m)$, где m — длина шаблона.

Поиск подстроки в тексте: $O(n)$, где n — длина текста.

Итоговая временная сложность: $O(n + m)$.

Оценка по памяти

$O(m)$ (для хранения префикс-функции, строк и вектора вхождений).

Вывод

В ходе выполнения лабораторной работы был изучен алгоритм Кнута-Морриса-Пратта. Была написана реализация данного алгоритма для решения задач. КМП – эффективный алгоритм, осуществляющий поиск подстроки в строке. Время работы алгоритма линейно зависит от объёма входных данных, то есть разработать асимптотически более эффективный алгоритм невозможно.