

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

**ОТЧЕТ**  
**по лабораторной работе №4**  
**по дисциплине «Построение и анализ алгоритмов»**  
**Тема: Алгоритм Кнута-Морриса-Пратта**

Студент гр. 7304

\_\_\_\_\_

Нгуен К.Х.

Преподаватель

\_\_\_\_\_

Филатов А.Ю.

Санкт-Петербург

2019

## Цель работы.

Исследование алгоритма Кнута-Морриса-Пратта, реализация программы для решения задач с использованием этого алгоритма

## Задание

### 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  $T$ .

Вход:

Первая строка -  $P$

Вторая строка -  $T$

Выход:

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

### 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.

Вход:

Первая строка -  $(A)$

Вторая строка -  $(B)$

Выход:

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

## Ход работы

## **Задание 1**

Алгоритм КМР, включающий 2 части: сначала мы должны создать таблицу КМР (или LSP, которая хранит длину самого длинного суффикса, который также является префиксом строки), затем мы используем эту таблицу, чтобы найти шаблон в строке за один проход

1. Написана функция `build_LSP_table`, которая принимает строку шаблона и создает таблицу LSP.

2. Написана функция КМР, которая принимает текст, таблицу, созданную функцией на последнем шаге, и строку шаблона, в результате получается вектор, в котором хранятся начальные точки каждого совпадения шаблона.

3. Написана основная функция для получения параметров и вывода результата в соответствии с заданием.

## **Задание 2**

Для второй задачи мы можем повторно использовать большую часть кода из первой задачи

Я добавил механизм, позволяющий указателю строки возвращаться к началу, когда он достигнет конца строки за один раз. Мы также можем сделать это по-другому: мы можем продублировать первую строку (например, от `abc` до `abcabc`), а затем найти в ней вторую строку, если она найдена и длины совпадают, то это цикл.

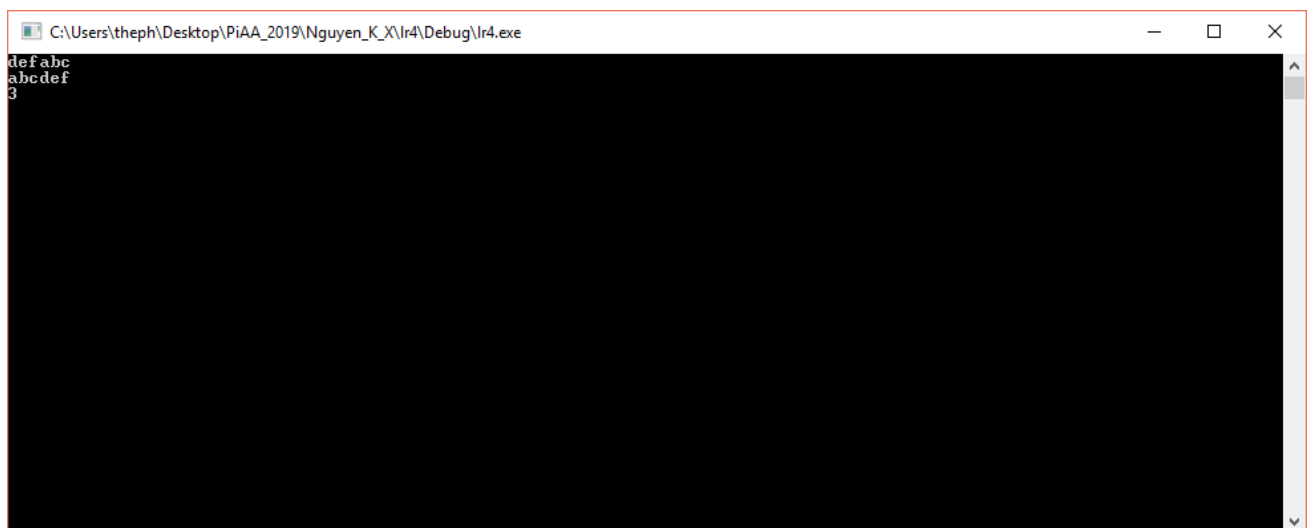
# Экспериментальные результаты.

## Задание 1



```
C:\Users\theph\Desktop\PiAA_2019\Nguyen_K_X\lr4\Debug\lr4.exe
ab
abab
0,2
```

## Задание 2



```
C:\Users\theph\Desktop\PiAA_2019\Nguyen_K_X\lr4\Debug\lr4.exe
defabc
abcdef
3
```

## Выводы.

В результате лабораторной работы я изучил алгоритм КМР для быстрого поиска в строке и использовал его для решения задач.

Алгоритм КМР особенно полезен для поиска шаблона в очень длинной строке, в которой наивный алгоритм поиска занимает много времени для поиска.