

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. 9383

Звега А.Р.

Преподаватель

Фирсов М.А.

Санкт-Петербург

2021

Цель работы.

Изучить алгоритм Кнута-Морриса-Пратта, для нахождения подстроки в строке и определения циклического сдвига.

Задание.

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 .

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.

Выполнение работы.

Алгоритм Кнута-Морриса-Пратта использует результаты префикс-функции от подстроки. На каждой итерации происходит проверка символов строки и подстроки. В случае несовпадения индекс подстроки перемещается на значение префикс функции для предыдущего символа и сравнивается с текущим символом строки.

Для поиска циклического сдвига первая строка удваивается и происходит поиск вхождения второй строки в первую, индекс на выходе которого является ответом для циклического сдвига

Анализ алгоритма.

Значения префикс функции высчитываются за линейное время $O(n)$, где n - длина подстроки. Обход основной строки выполняется один раз, откуда затраченное время тоже линейно и равно $O(m)$, где m - длина строки. Общее время составляет $O(n+m)$.

Таблица 1. Результаты тестирования алгоритма

Ввод	Вывод
123 123123123	0,3,6
aaa abbbbbbaabbbbfbbbaabbbbbbbfpffppdaa	-1
fadsfasdfsadfasdf asdfasdf	-1

Таблица 2. Результаты тестирования циклического алгоритма

Ввод	Вывод
123456789 789123456	6
aaaaaaaaaaa aaaaaaaaaab	-1
aaa a	-1

Выводы.

В ходе выполнения лабораторной работы был реализован алгоритм КнутаМорриса-Пратта. Был получен навык работы с префикс функцией. Так же был реализован алгоритм для поиска циклического сдвига.

ПРИЛОЖЕНИЕ А.

ИСХОДНЫЙ КОД.

Название файла lab4_1.py

```
def prefix(s):  
    l = len(s)  
    P = [0]*l  
    i, j = 0, 1  
    while j < l :  
        if s[i] == s[j]:  
            P[j] = i + 1  
            i += 1  
            j += 1  
        elif i:  
            i = P[i - 1]  
        else:  
            P[j] = 0  
            j += 1  
    return P  
  
def kmp(sub, text):  
    sub_len = len(sub)  
    text_len = len(text)  
    if not text_len or sub_len > text_len:  
        return []  
    P = prefix(sub)  
    entries = []
```

```

i = j = 0
while i < text_len and j < sub_len:
    if text[i] == sub[j]:
        if j == sub_len - 1:
            entries.append(i - sub_len + 1)
            j = 0
        else:
            j += 1
        i += 1
    elif j:
        j = P[j-1]
    else:
        i += 1
return entries

```

```

if __name__ == '__main__':
    sub = input()
    s = input()
    P = kmp(sub, s)
    if P:
        print(*P, sep=',')
    else:
        print(-1)

```