

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. 7304

Соколов И.Д

Преподаватель

Филатов А.Ю.

Санкт-Петербург

2019

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

Описание алгоритма.

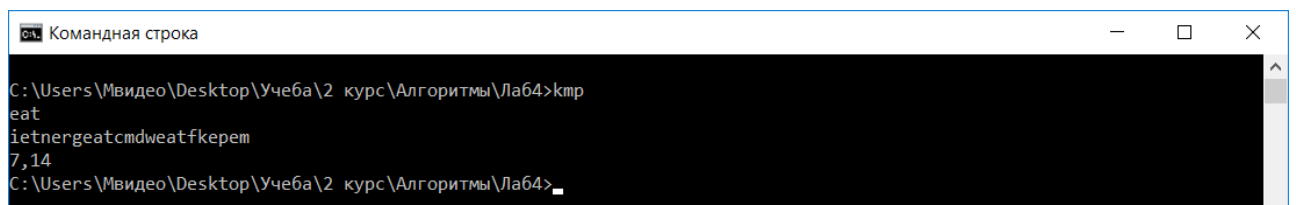
1. Для строки, вхождения которой мы будем искать, вычислим префикс функцию.
2. Каждый символ исходной строки в цикле сравнивается с элементом из строки, вхождения которой проверяется, по индексу j , который определяется как значение префикс функции по индексу $j-1$, если сравниваемые элементы не равны, и увеличивается, если элементы равны. Если j становится равным размеру подстроки, то в массив вхождений записывается индекс начала вхождения.
3. Индексы всех вхождений выводятся на экран. Если вхождения не были найдены, выводится -1.

Описание функций и структур данных.

Описание функций:

1. `vector<int> KMP(string init, string search)` – основная функция, выводящая индексы начал вхождений подстроки в строку;

Тестирование.



```
Командная строка
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>kmp
eat
ietnergeatcmdweatfkepem
7,14
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>
```

рис. 1

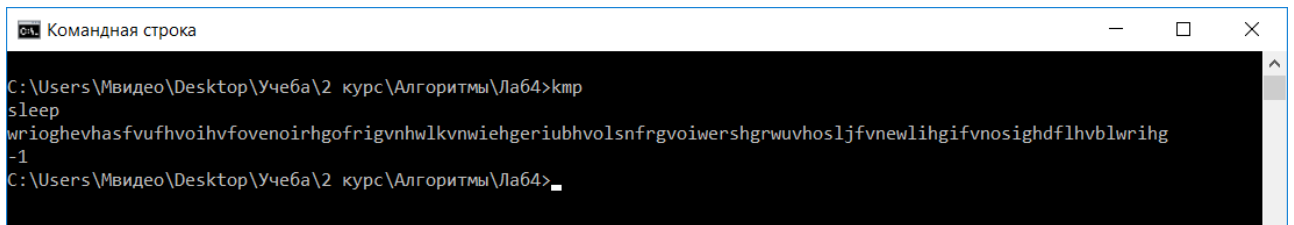


рис.2

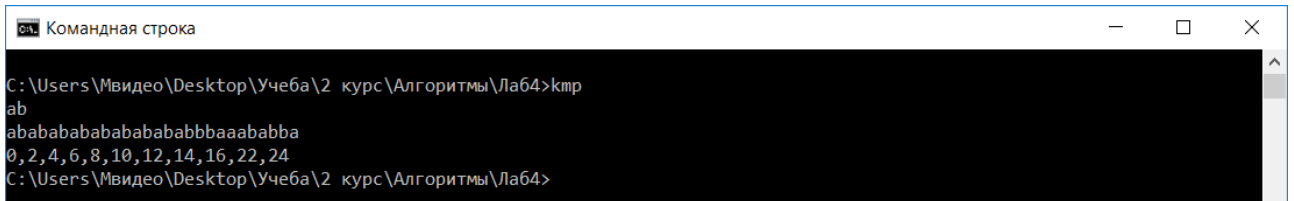


рис. 3

Задание 2.

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

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

Вход:

Первая строка - A

Вторая строка - B

Выход:

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

Описание алгоритма.

1. Разделили вторую строку пополам, поскольку хотя бы одна из ее половин должна полностью (без обрезов) содержаться в первой строке. Если число символов нечетное, то в первой половине содержится меньшая половина.
2. Выполняем поиск первой половины второй строки в первой строке по алгоритму Кнута-Морриса-Пратта, выводим индекс первого вхождения, либо -1 , если вхождение найдено не было.
3. Если вхождение не было найдено, выполняем поиск второй половины второй строки в первой строке, выводим индекс первого вхождения, либо

-1, если вхождение найдено не было. Если вхождение было найдено, сравниваем оставшиеся значения первой строки со второй половиной. Если все совпало – выводим индекс начала второй строки в первой. В противном случае выводим -1.

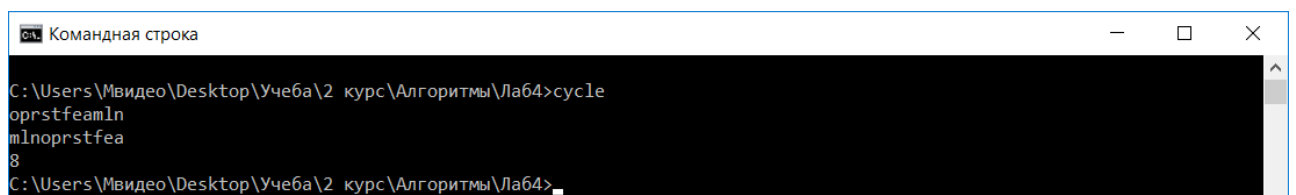
4. Если при вхождении второй половины совпадений найдено не было, выводим -1, иначе сравниваем оставшиеся значения первой строки с первой половиной. Если все совпало – выводим индекс начала второй строки в первой.

Описание функций и структур данных.

Описание функций:

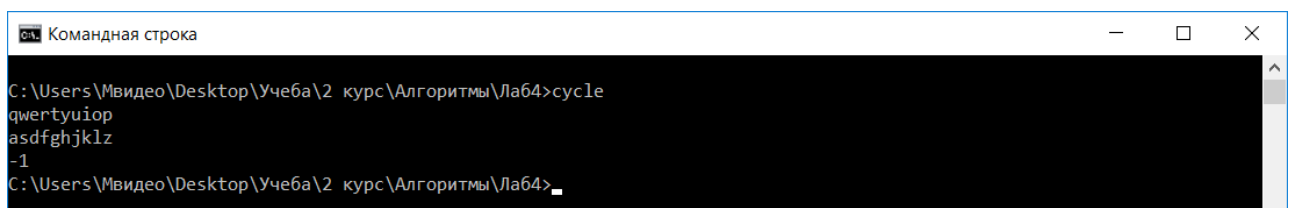
1. **int KMP(string init, string search)** – алгоритм Кнута-Морриса-Пратта, вычисляет и возвращает первый индекс вхождения подстроки в строку;
5. **int position (string first, string second, string f_half, string s_half, int index, int flag)** – функция, проверяющая совпадение остаточной части первой строки и непроверенной половины второй строки;
6. **int find_cycle (string first, string second)** – основная функция проверки на цикличность, возвращает индекс начала второй строки в первой;

Тестирование.



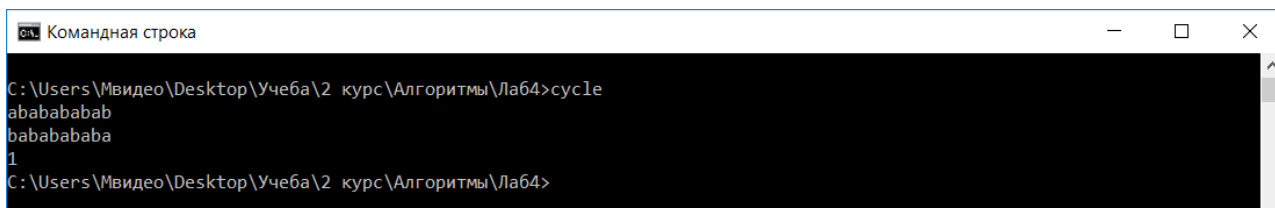
```
Командная строка
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>cycle
oprstfeamln
mlnoprstfea
8
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>
```

рис. 4



```
Командная строка
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>cycle
qwertyuiop
asdfghjklz
-1
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>
```

рис. 5



```
Командная строка
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>cycle
ababababab
bababababa
1
C:\Users\Мвидео\Desktop\Учеба\2 курс\Алгоритмы\Лаб4>
```

рис. 6

Выводы.

В ходе лабораторной работы по теме «Алгоритм Кнута-Морриса-Пратта». Был изучен и реализован на языке программирования C++ алгоритм Кнута-Морриса-Пратта для поиска подстроки в строке. Было изучено понятие префикс-функции, которая определяет длину наибольшего префикса подстроки, который совпадает с ее суффиксом. Так же был реализован алгоритм поиска циклического сдвига в строке.