**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»**

Студентка гр. 6381 \_\_\_\_\_\_\_\_\_\_\_\_\_ Лопатина А.С.

Преподаватель \_\_\_\_\_\_\_\_\_\_\_\_\_ Филатов А.Ю.

Санкт-Петербург

2018

**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.**

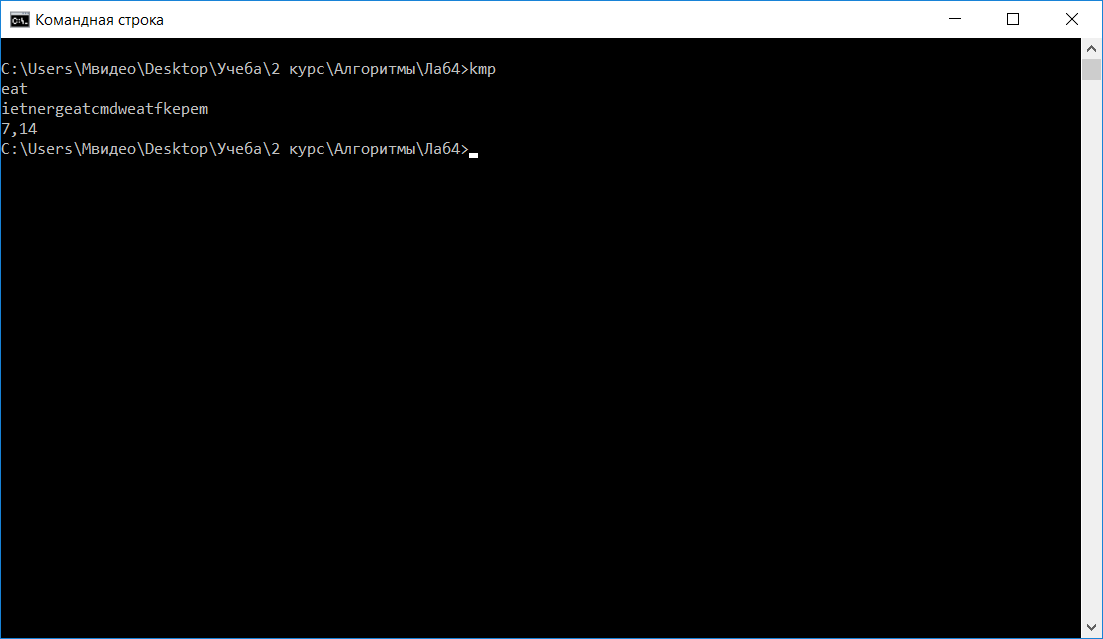


рис. 1

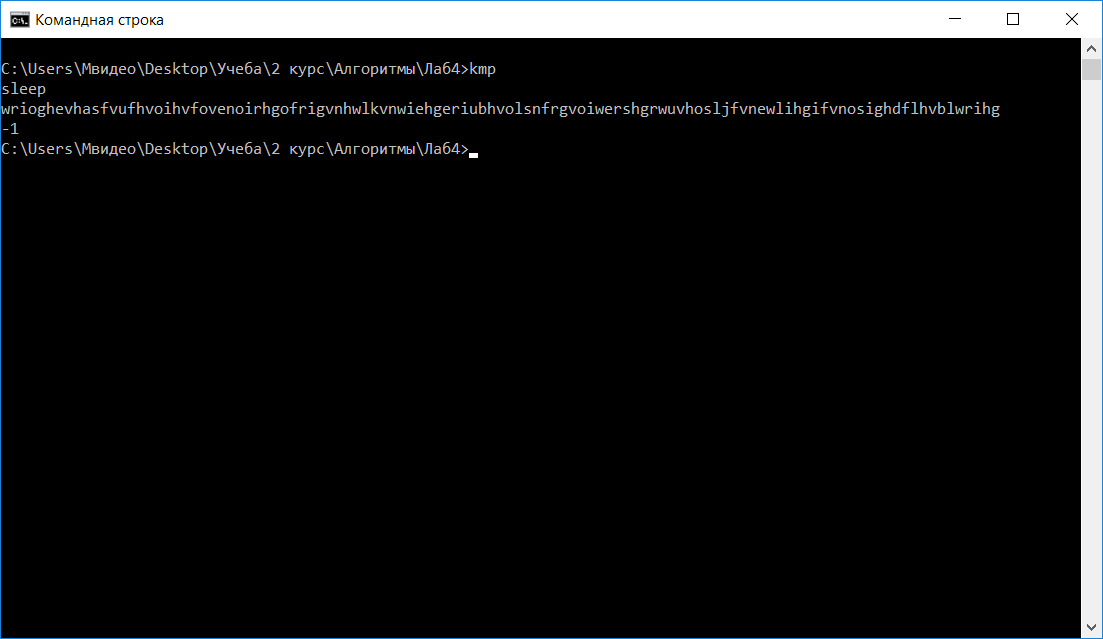


рис.2

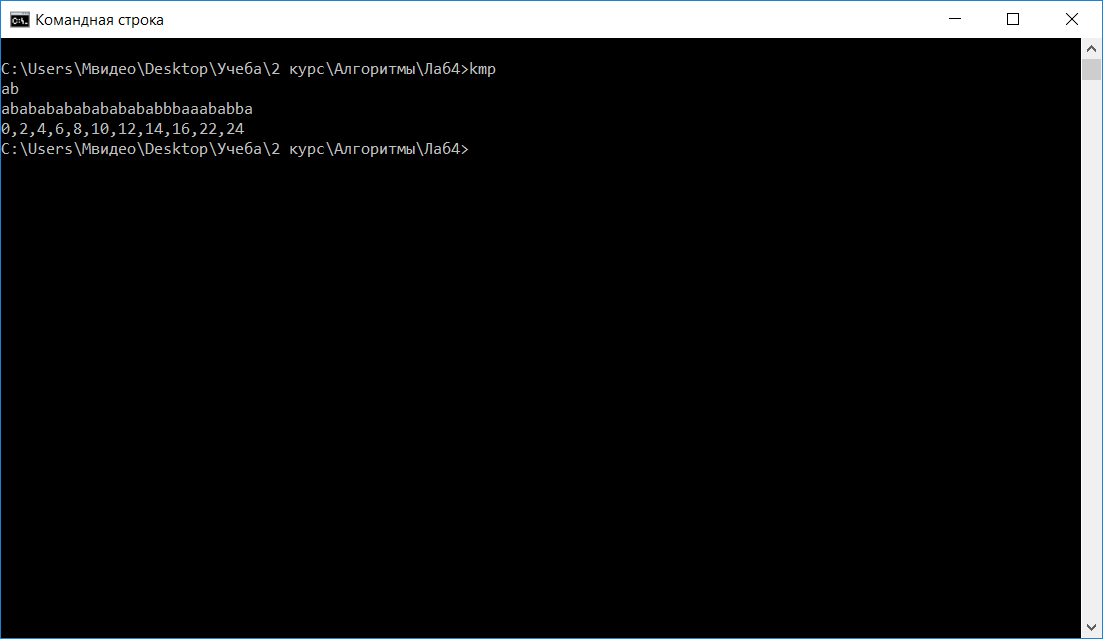


рис. 3

**Задание 2.**

Заданы две строки $A$ ($ |A| \leq 5000000 $) и $B$ ($ |B| \leq 5000000 $).  
Определить, является ли $А$ циклическим сдвигом $В$ (это значит, что $А$ и $В$ имеют одинаковую длину и $А$ состоит из суффикса $В$, склеенного с префиксом $В$). 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.

**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 вхождений 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;
2. **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;
3. **int find\_cycle (string first, string second) –** основная функция проверки на цикличность, возвращает индекс начала второй строки в первой;

**Тестирование.**

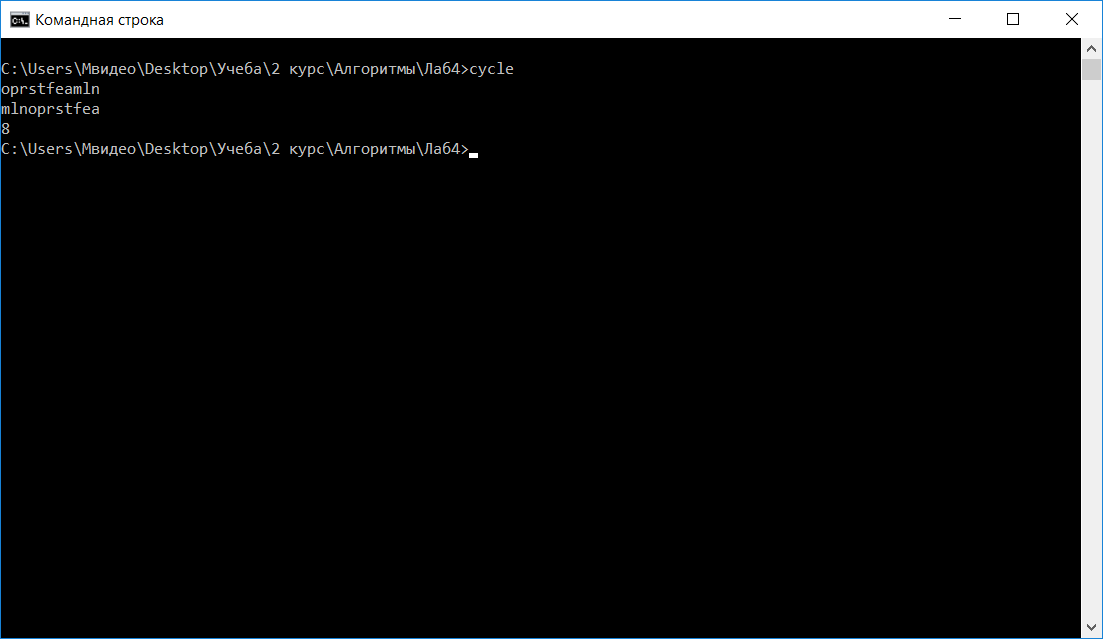


рис. 4

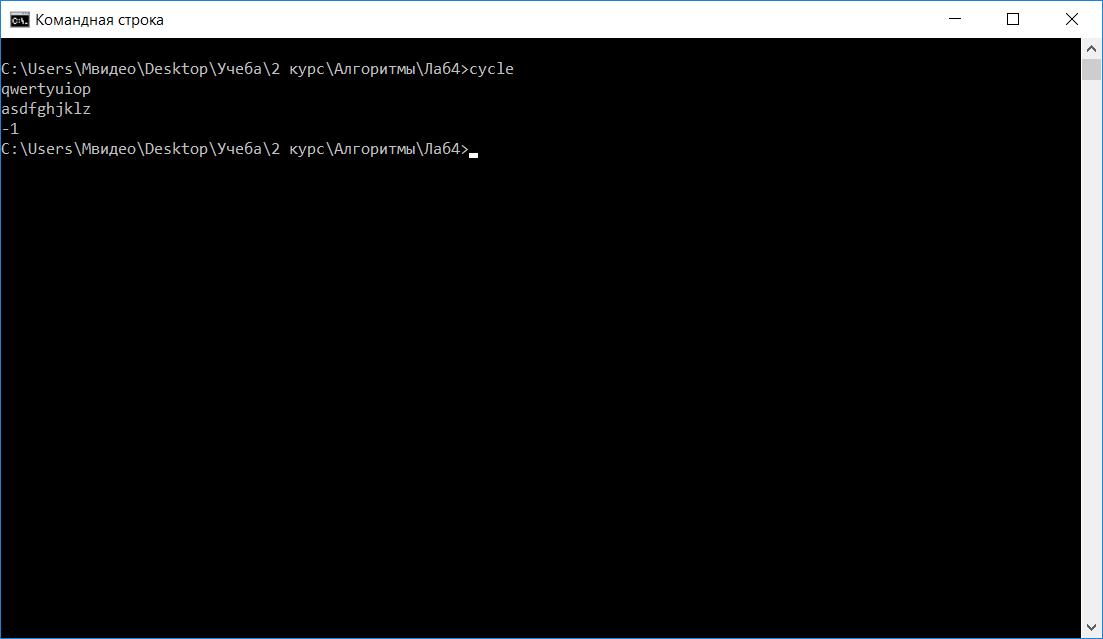


рис. 5

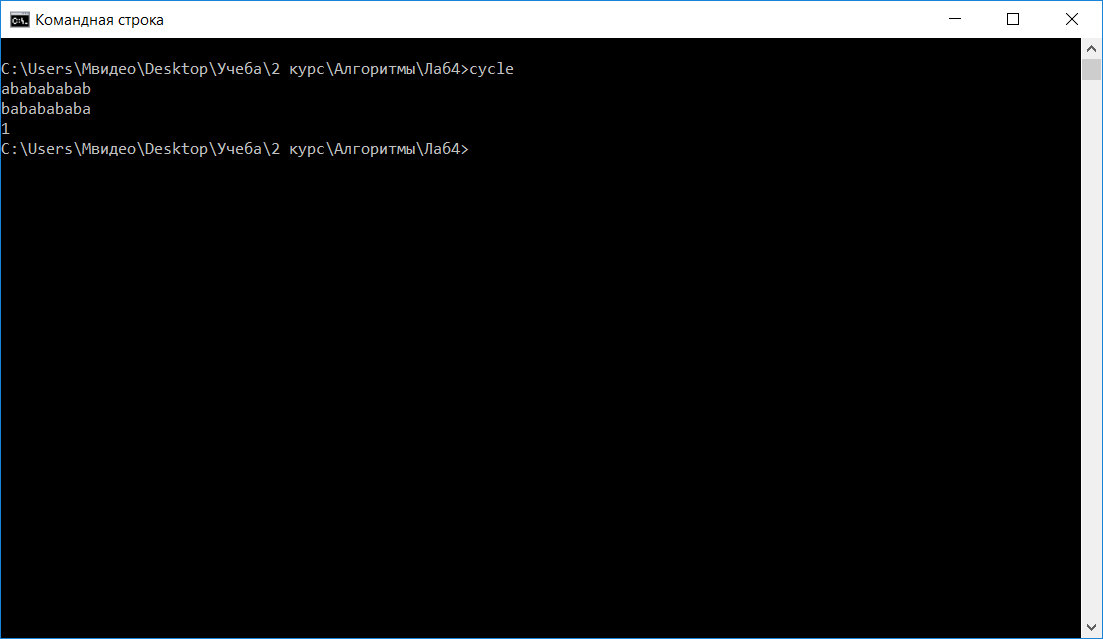


рис. 6

**Выводы.**

В ходе лабораторной работы мы подробно ознакомились с алгоритмом Кнута-Морриса-Пратта и изучили понятия префикс-функции, префикса, суффикса, собственного суффикса. В результате выполнения лабораторной работы были разработаны программы, которые вычисляют индексы вхождений подстроки в строку, а также проверяют, не является ли одна строка циклическим сдвигом второй.