**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: **Коммивояжер**

**Алгоритм Кнута-Морриса-Пратта**

|  |  |  |
| --- | --- | --- |
| Студентка гр. 1304 |  | Хорошкова А.С. |
| Преподаватель |  | Шевелева А.М. |

Санкт-Петербург

2023

**Цель работы.**

Изучение алгоритма Кнута-Морриса-Пратта для поиска подстроки в строке. Модификация алгоритма для установления размера сдвига одной строки относительно другой.

**Задание.**

1) Первое задание. Реализуйте алгоритм КМП и с его помощью для заданных шаблона P (∣P∣≤15000) и текста T (∣T∣≤5000000) найдите все вхождения P в T.

2) Второе задание. Заданы две строки A (∣А∣≤5000000) и B (∣B∣≤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). Если A является циклическим сдвигом B, индекс начала строки B в A, иначе вывести −1. Если возможно несколько сдвигов вывести первый индекс.

**Выполнение работы.**

**Описание алгоритма.**

1) Первое задание. Был использован алгоритм поиска подстроки Кнута-Морриса-Пратта. Был построен вектор префикс-функций на основе строки needle (игла) с помощью функции generate\_prefix\_function(). Далее производится поиск строки needle в строке haystack (стог сена) с помощью функции kmp(). Благодаря ранее найденной префик-функции алгоритм «перескакивает» заведомо совпавшие участки строк, продолжая поиск в месте с неисследованными буквами. Затем если найдены вхождения needle в haystack, то выводится все индексы этих вхождений, если вхождения не найдены, то выводится -1.

2) Второе задание. Был модифицирован алгоритм Кнута-Морриса-Пратта в методе kmp\_circle(). Внесены следующие изменения в классический алгоритм. Если длины строк needle и haystack не равны, то программа завершается с выводом -1 (строки не являются сдвигом друг друга). Также выводится только первое нахождение needle. Далее модифицированный алгоритм запускается на строке needle удвоенной строке haystack.

Оба алгоритма успешно прошли все тесты на платформе stepik. Результаты представлены на Рисунке 1.

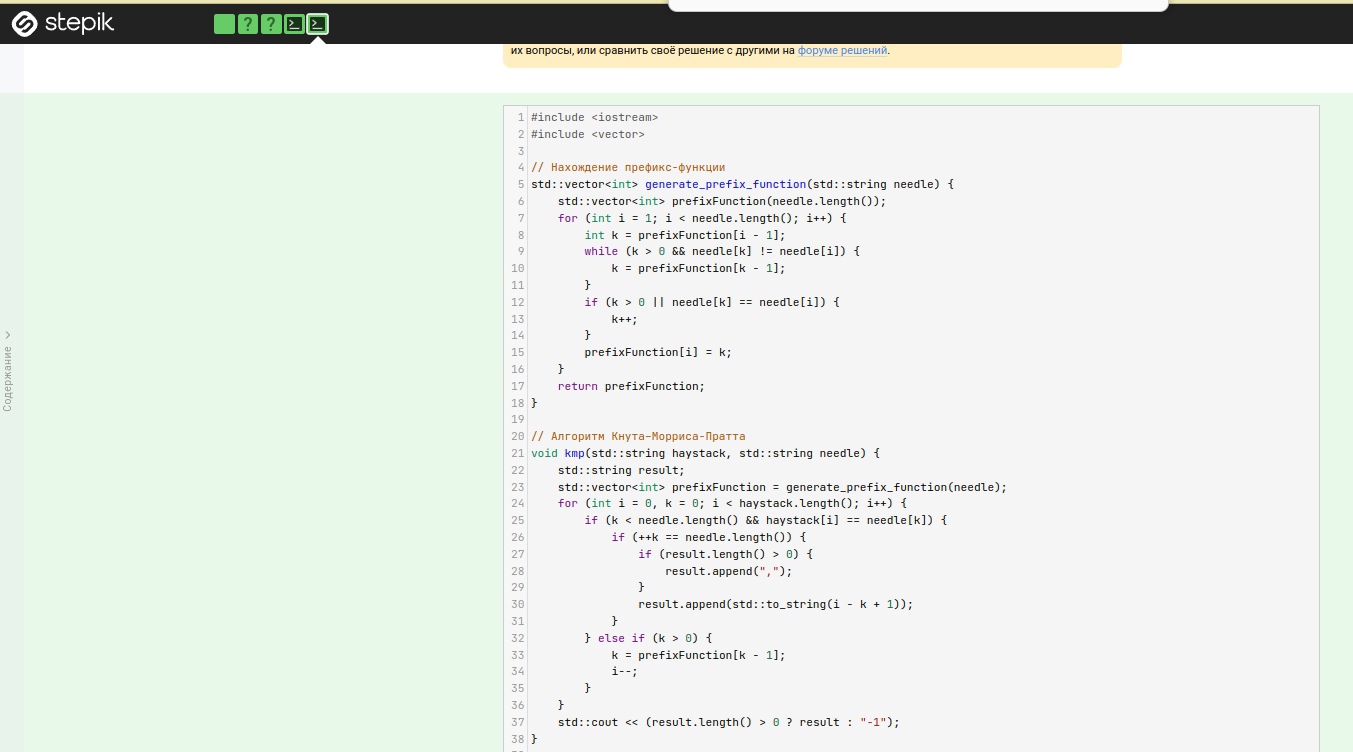


Рисунок - скриншот задания на платформе stepik

**Описание функций и структур данных.**

Для реализации работы двух алгоритмов были реализованы следующие функции:

std::vector<int> generate\_prefix\_function(std::string needle) — функция, принимающая на вход строки needle, и возвращающая на выходе вектор со значениями префикс-функции.

void kmp(std::string haystack, std::string needle) — функция, реализующая алгоритм Кнута-Морриса-Пратта. Она выводит на экран индексы букв, с которых начинаются все вхождения строки needle в строку haystack. Если вхождений не найдено, возвращает -1.

void kmp\_circle(std::string haystack, std::string needle) — функция, реализующая модификацию алгоритма Кнута-Морриса-Пратта для определения, являются ли строки циклическим сдвигом друг друга. функция выводит на экран индекс начала строки needle в строке haystack. Если haystack не является циклическим сдвигом needle, возвращается -1.

int main() — точка входа в программу. В методе main() считываются строки needle и haystack и запускается необходимая в данный момент функция (kmp или kmp\_circle).

**Выводы.**

В ходе лабораторной работы на практике был изучен алгоритм поиска подстроки в строке Кнута-Морриса-Пратта.

В рамках реализации первой задачи был написан классический алгоритм Кнута-Морриса-Пратта. Алгоритм находит первую строку во второй и выводит все индексы начала вхождения первой строки во вторую либо -1, если вхождения не найдены.

В рамках реализации второй задачи был модифицирован алгоритм Кнута-Морриса-Пратта таким образом, чтобы он определял, является ли одна строка циклическим сдвигом второй и выводит либо величину сдвига, либо -1, если строки не являются циклическим сдвигом друг друга.

Оба алгоритма успешно прошли все тесты на платформе stepik. Результаты представлены на Рисунке 1.