**Київський національний університет імена Тараса Шевченка  
Факультет комп’ютерних наук та кібернетики**

**Алгоритми та складність  
Лабораторна робота №6**

**Перевірка строки на зсув**

**Звіт**

**Підготував:**студент групи К-29  
Дацюк Віталій Олегович

**Київ-2019**

**1.Постановка завдання**

Розробіть алгоритм, який за лінійний час визначав би, чи є текстовий рядок Т циклічним зсувом іншого рядка Т\* (наприклад, *abc* та *cab*).

**2.Опис алгоритму**

Для вирішення цієї задачі можна використати алгоритм Кнута-Моріса-Пратта, який використовує префіксну функцію для знаходження підрядка в стрічці.

Для того, щоб зрозуміти, що і як робить префіксна функція, подивимося на рядок

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| a | **a** | **b** | a | **a** | **b** | a | **a** | **a** | a | b | a | a | b | а | a | a | b |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 0 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 2 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 3 |

Рядок під нею - номер (позиція) символу в рядку (для зручності опису алгоритму вважаємо початок масиву з номеру 1), а сама нижня строка - масив M довжин префіксів, ключ до розуміння префікс-функціі. Візьмемо символ з номером 7 (це a) і для K від 1 до 6 розглянемо рядки-префікси (підрядок, що починається з першого індексу рядка) і суфікси (підрядок, останній символ якої в рядку в позиції 7 (це «наш» символ a) довжини K.

Позначимо S-вихідний рядок, S (n) -початок (префікс) рядка S довжини n, S [n]-символ на позиції n рядка S. M [n] значення в масиві, S (M [n]) - той самий рядок, який є префіксом і суфіксом максимальної довжини для позиції n (для стислості позначимо її P (n)). Рядок P (n) як би «віртуальний», він не формується і нікуди не пишеться. Це просто початковий фрагмент1 початкового рядка S довжини M [n]. І цей початковий фрагмент1 збігається (як послідовність символів) з фрагментом2 довжини M [n], останній символ якого знаходиться в позиції n. Якщо M [n] = 0, то збігу немає .Маємо: позиція 7 масиву заповнена значенням М [7] = 4, найдовший рядок P (7) = 'aaba' довжини 4,треба перейти до позиції 8 і заповнити M [8]. Можна просто порахувати всі префікси і суфікси довжиною від 1 до 7, порівняти їх і записати максимальну довжину в позицію 8. Але ми підемо іншим шляхом (слідом за **АЛГОРИТМОМ КНУТА-МОРІСА-ПРАТТА).** Нехай знайдено максимально довгий рядок P (8) довжини k, який є префіксом і суфіксом для позиції 8. Рядок p7 з перших k-1 символів є префіксом і суфіксом для позиції k-1. Не факт, що для 7ї позиції він найдовший. Однак, якщо виявилося, що p7 = P7, то P8 - це розширення P7 на один символ. Щоб перевірити, чи можна розширити P7 на одну позицію, треба перевірити, чи збігається символ що додається в суфікс (це символ S [8] = a) з наступним символом префікса. Наступний символ префікса a знаходиться в позиції М [7] + 1 = 5. Якщо збігся (а в нашому випадку він збігся), то завдання виконане - М [8] = М [7] +1, а P (8) = P (7) + символ з 8 позиції S [8] = a. Отримуємо P (8) = 'aabaa'. При успішному розширенні треба всього одне порівняння для обчислення чергового значення масиву.

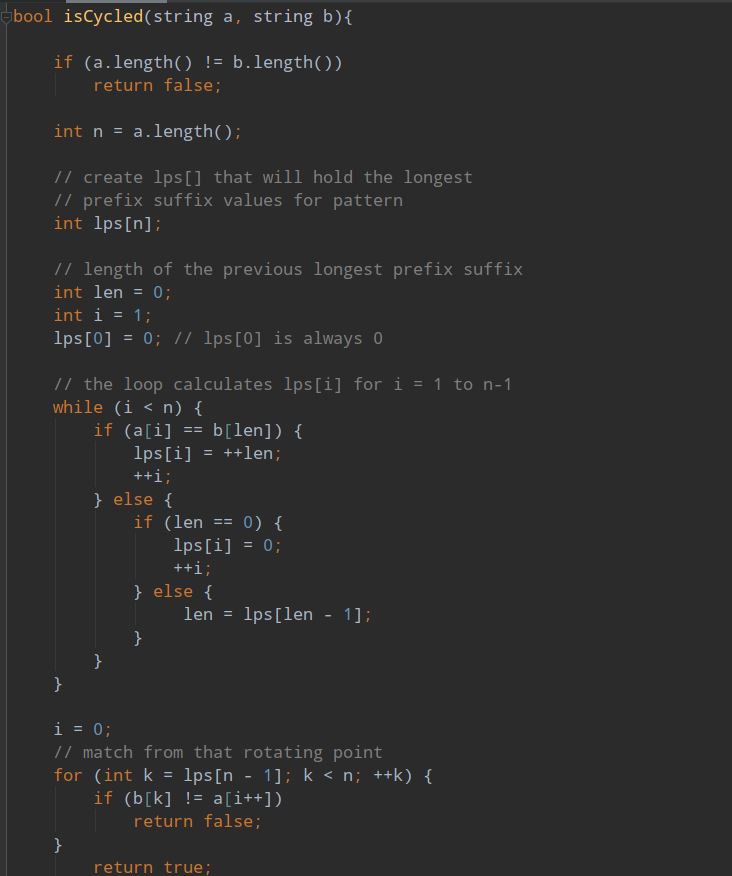
Тепер інший випадок - P8 розширити не вдалося, тобто символ S [9] = a не збігся із символом рядка S у позиції M [8] + 1 = 6 b. Суфікс розширюється легко (оскільки новий символ просто дописується в кінець), а з префіксом проблеми, тому символ що додається в суфікс може не збігтися з черговим символом префікса. Якщо префікс P (k) не підходить, треба знайти інший такий префікс, коротший, у якого такий же суфікс і спробувати розширити його. Але префікс коротший, причому з таким же суфіксом - це S [M [M [k]]), тому що при заповненні масиву М кожен елемент містить довжину максимально довгого префікса з таким же суфіксом. Тому, якщо не вдалося розширити S (M [k]) пробуємо так само розширити S (M [M [k]]) і т.д, поки не співпаде або поки не отримаємо нульову довжину.Цикл перебору походящих префіксів закінчується дуже швидко, тому, що вся необхідна для цього інформація вже сидить в масиві М. Для нашого рядка P (8) - це просто розширення P (7) на один символ, де знадобилося 1 порівняння. Однак P (8) не вдалося розширити до P (9), оскільки S [9] = a, а S [M [8] + 1 = 6] = b. Раз не підійшов префікс P8 довжини M [8] = 5, пробуємо префікс довжини M [5] = 2. Він теж не підходить: S [2 + 1] = b. Пробуємо префікс довжини M [2] = 1 і його можна розширити, тому що S [1 + 1] = a. Тому M [9] = 2, на одиницю більше довжини розширюваного префікса. Для заповнення M [10] треба 2 порівняння .А ось щоб заповнити елементи з 11 по 17, знадобиться по одному порівнянні. В результаті розрахунок значень масиву займає час порядку О (довжина рядка).

Далі знаходимо найдовший спільний підрядок,який є одночасно префіксом другого і суфіксом першого рядків,на основі КМП. Це буде відправним пунктом подальшого порівняння. Якщо з цієї точки усі символи співпадуть в обох рядках, тоді рядок А є циклічним зсувом рядка Б.

Складність: **O(n)**

**3.Основні модулі програми**

* **Функція перевірки двох рядків на циклічний зсув**

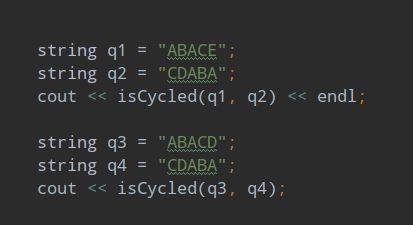


**4.Інтерфейс користувача**

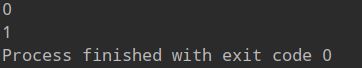
Для інтерфейсу користувача буде достатньо основних бібліотек с++, а саме <string>, адже уся робота виконана в одній функції на базі string.

**5.Тестові приклади**

Приклади:

****

Вивід на екран(1 – циклічний зсув, 0 – ні):

****

**Література:**

1. <https://www.geeksforgeeks.org/check-strings-rotations-not-set-2/>
2. <https://habr.com/ru/post/307220/>