*Государственное образовательное учреждение высшего профессионального образования*

**«Московский государственный технический университет имени Н. Э. Баумана»**

**(МГТУ им. Н.Э. Баумана)**

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

**ЛАБОРАТОРНАЯ РАБОТА №1**

**Отчет на тему:**

Расстояние Левенштейна

Янова Д. Ю.

ИУ7-53

Москва 2017

**1 Расстояние Левенштейна**

* 1. **Постановка задачи**

Реализовать алгоритм поиска расстояния Левенштейна, использую три алгоритма:

1. базовый
2. модифицированный
3. базовый через рекурсию
   1. **Алгоритм**

Пусть 𝑆1 и 𝑆2 — две строки (длиной 𝑀 и 𝑁 соответственно) над некоторым алфавитом, тогда редакционное расстояние (расстояние Левенштейна) 𝑑(𝑆1, 𝑆2) можно подсчитать по следующей рекуррентной формуле :

, (1.2)

, (1.3)

Где — разрешенные операции:

* 1. Замена символа (𝑅, 𝑟𝑒𝑝𝑙𝑎𝑐𝑒) Штраф 1.
  2. Вставка символа (𝐼, 𝑖𝑛𝑠𝑒𝑟𝑡) Штраф 1.
  3. Удаление символа (𝐷, 𝑑𝑒𝑙𝑒𝑡𝑒) Штраф 1.
  4. Совпадение символа (𝑀, 𝑚𝑎𝑡𝑐ℎ) Штраф 0.

В модифицированном алгоритме добавлена еще одна операция:

5. Перестановка символа (𝑋, 𝑒𝑥𝑐ℎ𝑎𝑛𝑔𝑒) Штраф 1.

В рекуррентную формулу добавляется еще один член минимума:

(1.4)

Алгоритм можно реализовать с помощью матрицы, двигаясь построчно или по столбцам, рассматривая «квадрат» значений: