|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ **Информатика и системы управления**

КАФЕДРА **«Программное обеспечение ЭВМ и информационные технологии» (ИУ7)**

НАПРАВЛЕНИЕ ПОДГОТОВКИ **09.03.01 Информатика и вычислительная техника**

**Отчет**

|  |  |
| --- | --- |
| **по лабораторной работе №** | 1 |

**Название:** Расстояние Левенштейна и Дамерау-Левенштейна

**Дисциплина:** Анализ алгоритмов

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Студент | ИУ7-52Б |  |  | В.А. Иванов |
|  | (Группа) |  | (Подпись, дата) | (И.О. Фамилия) |
|  |  |  |  |  |
| Преподаватель |  |  |  |  |
|  |  |  | (Подпись, дата) | (И.О. Фамилия) |

Москва, 2020

Оглавление

[**Введение** 3](#_Toc50489512)

[**1)** **Аналитическая часть** 4](#_Toc50489513)

[**2)** **Конструкторская часть** 5](#_Toc50489514)

[**3)** **Техническая часть** 6](#_Toc50489515)

[**4)** **Исследовательская часть** 7](#_Toc50489516)

[**Заключение** 8](#_Toc50489517)

[**Список литературы** 9](#_Toc50489518)

**Введение**

Расстояние Левенштейна – минимальное количество редакционных операций, которые необходимы для превращения одной строки в другую. Редакционными операциями в данном случае являются:

* Вставка символа
* Удаление символа
* Замена символа

Расстояние Дамерау-Левенштейна также учитывает и операцию транспозиции – перестановки двух соседних символов местами.

Данные расстояния имеют большое количество применений. Они используются для автокоррекции при выполнении поисковых запросов и печати на клавиатуре, а также в биоинформатике для сравнения генов, представленных в строковом формате.

1. **Аналитическая часть**

Целью лабораторной работы является реализация и сравнение алгоритмов поиска расстояний Левенштейна и Дамерау-Левенштейна.

Задачами лабораторной работы являются:

* Математическое описание расстояний Левенштейна и Дамерау-Левенштейна.
* Описание и реализация алгоритмов поиска расстояний.
* Проведение замеров процессорного времени работы алгоритмов при различных размерах строк. Оценка наибольшей используемой каждым алгоритмом памяти
* На основании экспериментов провести сравнительный анализ алгоритмов

Задача по поиску расстояний заключается в нахождении такого взаимного выравнивания строк и операций, которые будут иметь минимальный суммарный штраф. Штраф операций:

* Вставка (I) – 1
* Замена (R) – 1
* Удаление (D) – 1
* Совпадение (М) – 0
* Транспозиция (T) – 1

Для решения данной проблемы используется рекурентная формула вычисления расстояний. Пусть D(s1[1..i], s2[1..j]) – расстояние Левенштейна для подстроки s1 длиной i и s2 длиной j. Тогда D вычисляется как:

Аналогично рекурсивно представляется формула расстояния Дамерау-Левенштейна:

Последнее D учитывается в формуле при выполнении:

1. **Конструкторская часть**

Компьютерная фаф

1. **Техническая часть**

Компьютерная фаф

1. **Исследовательская часть**

Компьютерная фаф

**Заключение**

Компьютерная фаф

**Список литературы**

Компьютерная фаф