Министерство образования Республики Беларусь

Учреждение образования

«Брестский государственный технический университет»

Кафедра ИИТ

Лабораторная работа №3

По дисциплине: «Естественно-языковой интерфейс ИС»

# Тема: «Реализация алгоритмов автоматического морфологического и лексико-грамматического анализа текста естественного языка»

Выполнил:

Студент 3 курса

Группы ИИ-21

Карагодин Д. Л.

Проверила:

Якимук А. В.

Брест 2024

**Цель:** закрепить знания морфологического и лексико-грамматического анализа текста.

**Ход работы:**

|  |  |  |
| --- | --- | --- |
| **Номер**  **варианта** | **Язык текста** | **Формат документа** |
| 1 | Русский | Html |

**Задание 1.** Создать программную реализацию решения задачи о редакционном расстоянии методом динамического программирования между словами, подаваемыми на вход системы и словами текста.

**Задание 2.** Спроектируйте внешний интерфейс автоматизированной системы осуществляющей решения задачи о редакционном расстоянии между словами, подаваемыми на вход системы и словами текста.

**Задание 3.** Разработайте алгоритмы сервисных функций, предоставляемых Вашей системой.

**Задание 4.** Реализуйте разработанные алгоритмы в виде программ на языке программирования.

**Задание 5.** Протестируйте разработанную систему, используя различные входные тексты.

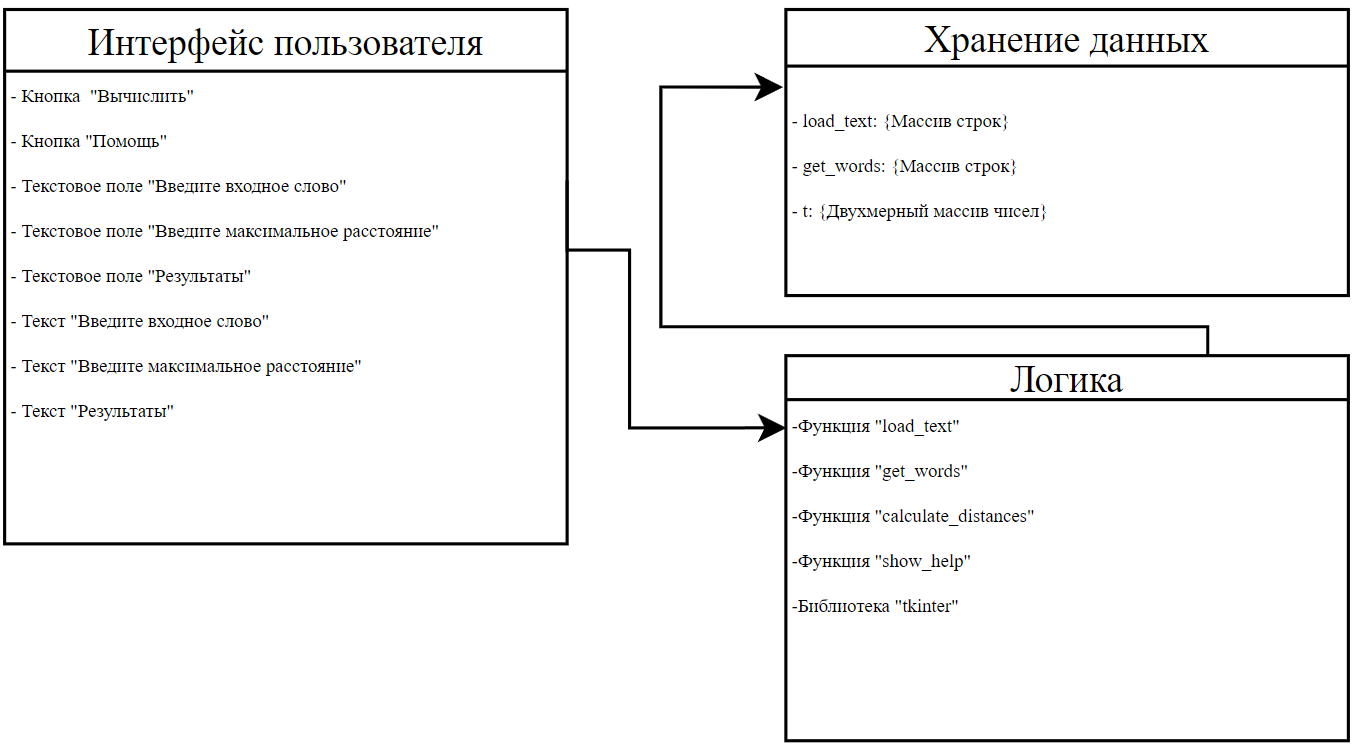
**Задание 6.** Разработайте и внедрите в систему средства помощи пользователю.

**Задание 7.** Протестируйте систему, включающую средства помощи пользователю.

**Методические указания:**

* на входе – слова естественно-языкового текста содержащие орфографические ошибки;
* на выходе – перечень слов, расстояние до которых от входного слова не превышает некоторого заданного Вами значения, упорядоченный по убыванию величины расстояния редактирования;
* интерфейс системы должен быть предельно простым и доступным для пользователей любого уровня, содержать понятный набор инструментов и средств, а также help-средства;
* наличие средств распечатки полученной на выходе информации.

**Структурно-функциональная схема:**



**Листинг:**

import tkinter as tk

# Задание 1: Функция вычисления редакционного расстояния

def edit\_distance(str1, str2):

    m = len(str1) + 1

    n = len(str2) + 1

    t = [[i + j for j in range(n)] for i in range(m)]

    for i in range(1, m):

        c = i - 1

        for j in range(1, n):

            d = j - 1

            t[i][j] = min(t[c][j] + 1, t[i][d] + 1, t[c][d] + (str1[c] != str2[d]))

    return t[m - 1][n - 1]

# Задание 2: Внешний интерфейс GUI

root = tk.Tk()

root.title("Редакционное расстояние")

# Задание 3: Сервисные функции

def load\_text(filename):

    with open(filename, "r", encoding="utf-8") as f:

        text = f.read()

    return text

def get\_words(text):

    words = text.split()

    return words

# Задание 4: Реализация сервисных функций

text = load\_text("index.html")

words = get\_words(text)

# Задание 5: Тестирование

def calculate\_distances():

    input\_word = input\_entry.get()

    max\_distance = int(max\_distance\_entry.get())

    results = []

    for word in words:

        distance = edit\_distance(input\_word, word)

        if distance <= max\_distance:

            results.append((word, distance))

    results.sort(key=lambda x: x[1], reverse=True)

    output\_text.delete("1.0", tk.END)

    for word, distance in results:

        output\_text.insert(tk.END, f"{word} ({distance})\n")

# Задание 6: Средства помощи пользователю

def show\_help():

    help\_window = tk.Toplevel(root)

    help\_window.title("Помощь")

    help\_text = tk.Text(help\_window)

    help\_text.insert(tk.END, "Помощь\n\nВведите входное слово:\nВведите максимальное расстояние:\nПолучите перечень слов с расстоянием редактирования меньше или равным максимальному.")

    help\_text.config(state="disabled")

    help\_text.pack()

# Интерфейс

input\_label = tk.Label(root, text="Введите входное слово:")

input\_entry = tk.Entry(root)

max\_distance\_label = tk.Label(root, text="Введите максимальное расстояние:")

max\_distance\_entry = tk.Entry(root)

calculate\_button = tk.Button(root, text="Вычислить", command=calculate\_distances)

help\_button = tk.Button(root, text="Помощь", command=show\_help)

output\_label = tk.Label(root, text="Результаты:")

output\_text = tk.Text(root, height=10)

input\_label.grid(row=0, column=0)

input\_entry.grid(row=0, column=1)

max\_distance\_label.grid(row=1, column=0)

max\_distance\_entry.grid(row=1, column=1)

calculate\_button.grid(row=2, column=0, columnspan=2)

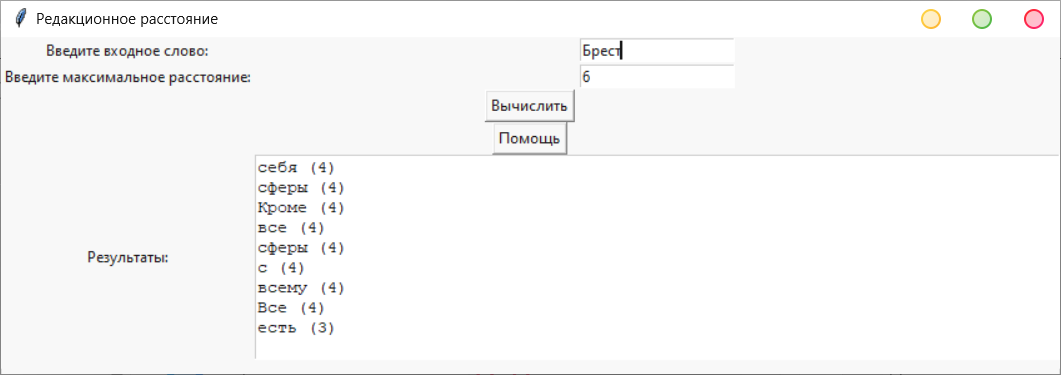
help\_button.grid(row=3, column=0, columnspan=2)

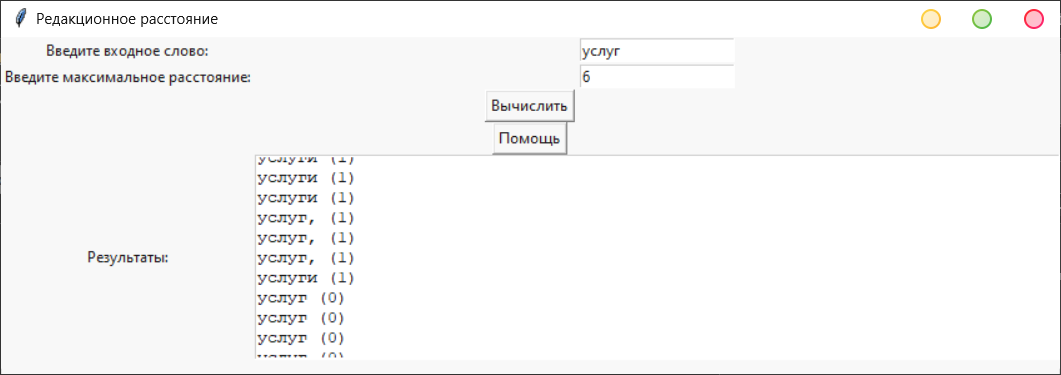
output\_label.grid(row=4, column=0)

output\_text.grid(row=4, column=1)

root.mainloop()

# Результат:





**Вывод:** в ходе выполнения лабораторной работы закрепил знания морфологического и лексико-грамматического анализа текста.