

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»

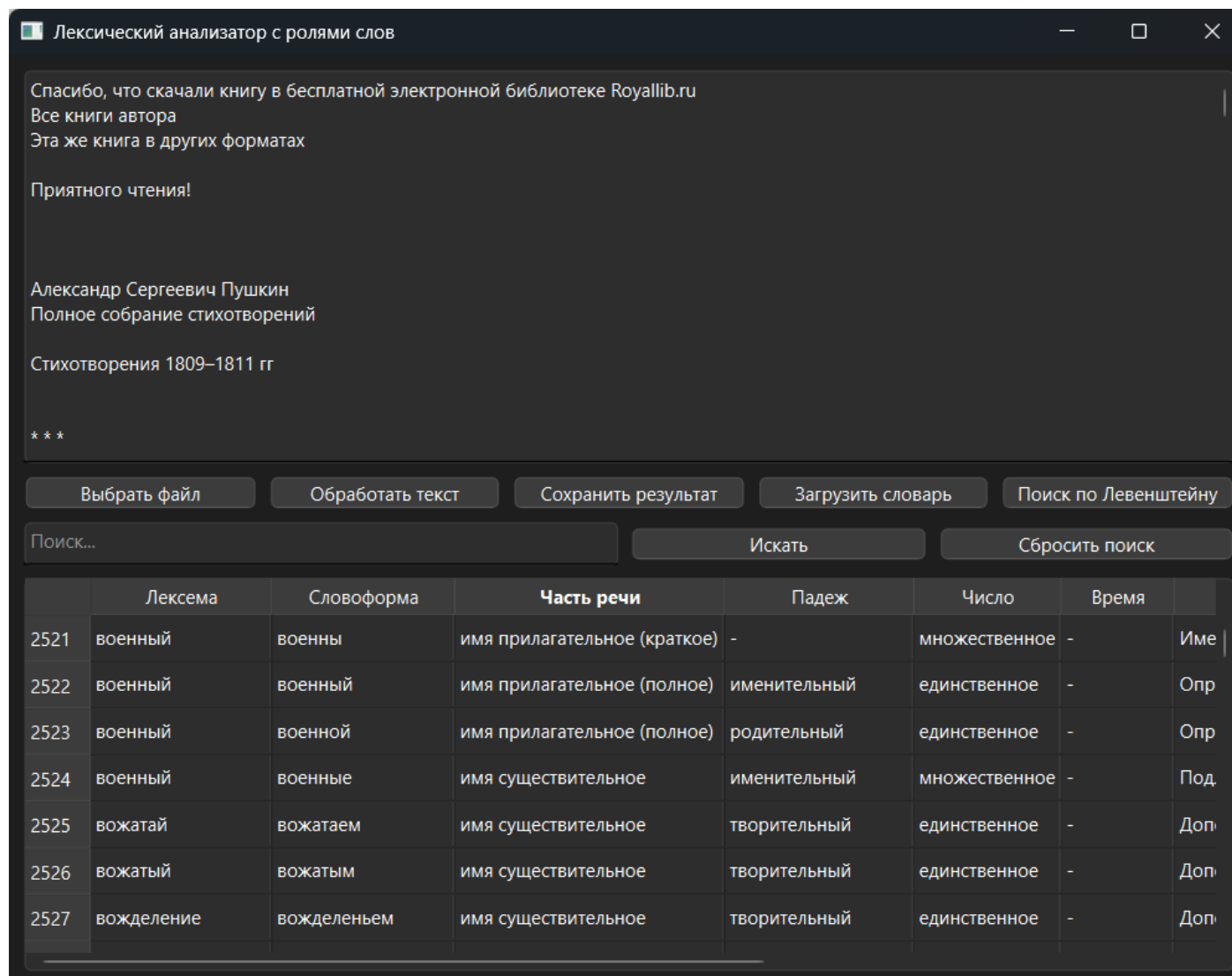
Выполнил:
Студент 3 курса
Группы ИИ-23
Романюк А. П.

Проверил:
Булей Е. В.

Ход работы

Задание:

1. *Создание программной реализации решения задачи о редакционном расстоянии*
Необходимо разработать программу, которая будет решать задачу о редакционном расстоянии между словами. Редакционное расстояние между двумя строками определяется как минимальное количество операций вставки, удаления и замены символов, необходимых для преобразования одной строки в другую.
2. *Проектирование внешнего интерфейса автоматизированной системы*
Необходимо спроектировать внешний интерфейс автоматизированной системы, которая будет осуществлять решение задачи о редакционном расстоянии.



Приложение с загруженным фалом docx

| | Лексема | Словоформа | Часть речи | Падеж | Число | Время | Роль в предложении | Частота |
|---|----------|------------|---------------------|--------------|---------------|-------|--------------------|-----------|
| 1 | а | а | союз | - | - | - | Союз | 339/11201 |
| 2 | абамелек | абамелек | имя существительное | именительный | единственное | - | Подлежащее | 1/11201 |
| 3 | август | августом | имя существительное | творительный | единственное | - | Дополнение | 1/11201 |
| 4 | август | августу | имя существительное | дательный | единственное | - | Дополнение | 1/11201 |
| 5 | август | августов | имя существительное | родительный | множественное | - | Дополнение | 1/11201 |
| 6 | авдей | авдей | имя существительное | именительный | единственное | - | Подлежащее | 3/11201 |
| 7 | авель | авеля | имя существительное | родительный | единственное | - | Дополнение | 1/11201 |

Характеристики выбранного слова

Поиск по редакционному расстоянию

Целевое слово:
сало

Максимальное расстояние:
2

Найти

| | Слово | Расстояние |
|---|-------|------------|
| 1 | сафо | 1 |
| 2 | пало | 1 |
| 3 | мало | 1 |
| 4 | стало | 1 |

Расстояние Левенштейна между введённым словом и всеми в словаре

Код программы:

```
from PySide6.QtWidgets import QDialog, QVBoxLayout, QLineEdit, QSpinBox, QTableWidgetItem,
QPushButton, QLabel, \
    QTableWidgetItem

from text_processor import TextProcessor

class LevenshteinSearchDialog(QDialog):
    def __init__(self, parent=None, words=None):
        super().__init__(parent)
        self.words = words
        self.setWindowTitle("Поиск по редакционному расстоянию")
        self.setGeometry(200, 200, 400, 300)

        layout = QVBoxLayout()

        self.target_word_edit = QLineEdit()
        self.max_distance_spin = QSpinBox()
        self.max_distance_spin.setRange(0, 10)
        self.max_distance_spin.setValue(2)
```

```

self.results_table = QTableWidgetItem()
self.results_table.setColumnCount(2)
self.results_table.setHorizontalHeaderLabels(["Слово", "Расстояние"])

search_button = QPushButton("Найти")
search_button.clicked.connect(self.perform_search)

layout.addWidget(QLabel("Целевое слово:"))
layout.addWidget(self.target_word_edit)
layout.addWidget(QLabel("Максимальное расстояние:"))
layout.addWidget(self.max_distance_spin)
layout.addWidget(search_button)
layout.addWidget(self.results_table)

self.setLayout(layout)

def perform_search(self):
    target = self.target_word_edit.text().lower()
    max_dist = self.max_distance_spin.value()

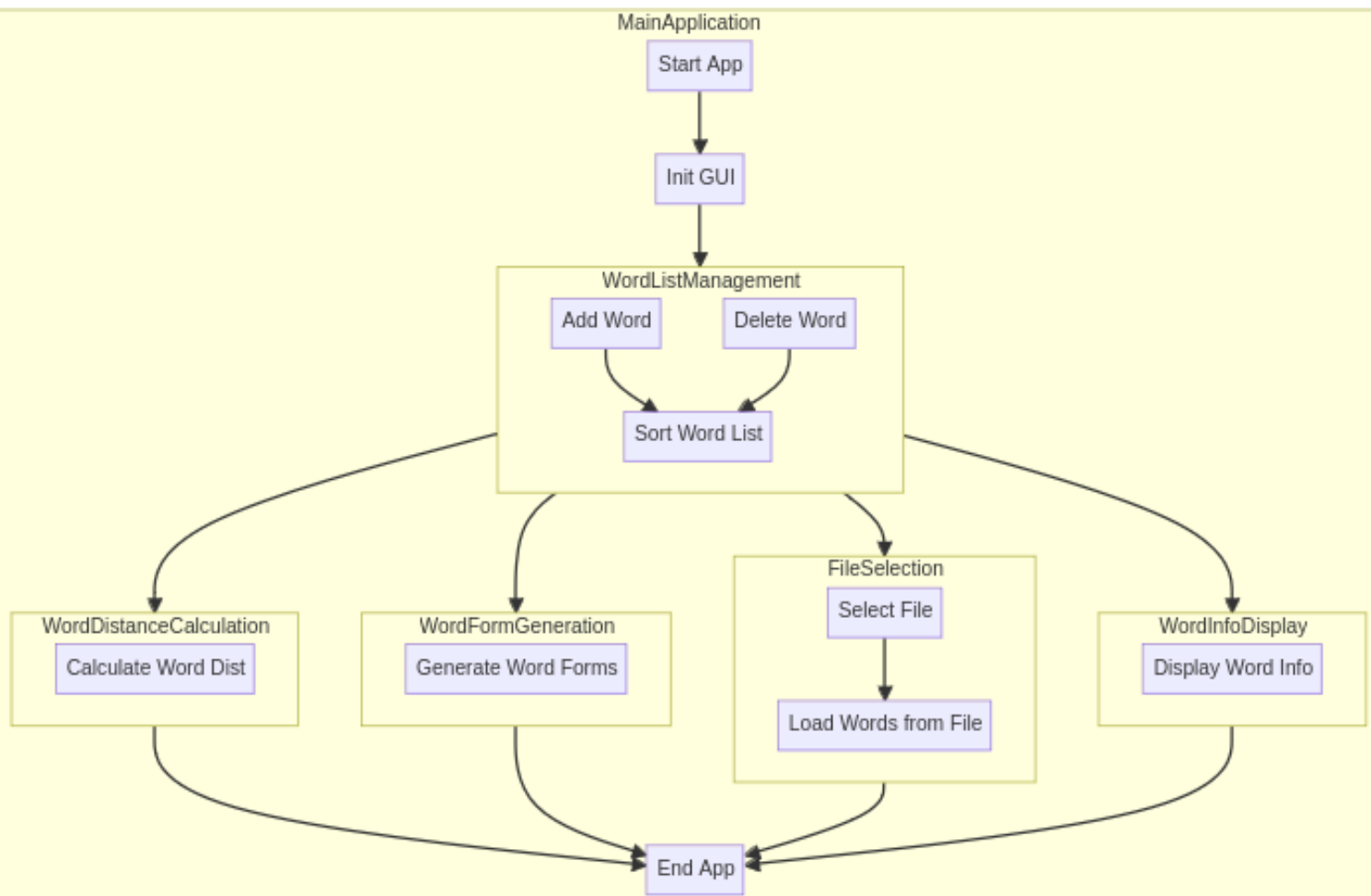
    self.results_table.setRowCount(0)

    if not target:
        return

    unique_words = list(set(self.words)) # Убираем дубликаты
    for word in unique_words:
        distance = TextProcessor.levenshtein_distance(target, word.lower())
        if distance <= max_dist:
            row = self.results_table.rowCount()
            self.results_table.insertRow(row)
            self.results_table.setItem(row, 0, QTableWidgetItem(word))
            self.results_table.setItem(row, 1, QTableWidgetItem(str(distance)))

    self.results_table.sortItems(1) # Сортировка по расстоянию
    self.results_table.resizeColumnsToContents()

```



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