

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования

«Южно-Уральский государственный университет
(национальный исследовательский университет)»
Высшая школа электроники и компьютерных наук
Кафедра «Электронные вычислительные машины»

РАЗРАБОТКА ПРОГРАММЫ ДЛЯ ПОИСКА ЗАИМСТВОВАНИЙ ПО ФАЙЛАМ ЭЛЕКТРОННЫХ ТАБЛИЦ

Руководитель работы,
к.пед.н., доцент каф. ЭВМ
М.А. Алтухова

Автор работы,
студент группы КЭ-405
Е.Е. Иванов

Содержание

Актуальность работы.....	3
Цель и задачи.....	4
Обзор аналогов.....	5
Анализ имеющихся решений.....	6
Косинусное сходство.....	7
Функциональные требования к программе.....	8
Нефункциональные требования к программе.....	9
Проектирование функциональной модели программы.....	10
Блок-схема алгоритма работы программы.....	11
Блок-схема алгоритма функции установления маркеров ошибок.....	13
Выбор средств реализации.....	14
Реализация.....	15
Заключение.....	18
Библиографический список.....	19

Актуальность работы

В настоящее время с развитием возможностей интернета обостряется такая проблема, как проблема плагиата. В Интернете существует огромное множество работ, статей, материалов, и поэтому найти что-то уникальное становится достаточно проблематично. Так же, эта проблема сказалась и на образовании. Студенты и школьники зачастую выбирают наиболее простой способ выполнения некоторых работ, а именно заимствовать у товарищей или найти готовые решения в Интернете.

Цель и задачи

Цель: разработать программу, которая будет проверять файлы электронных таблиц на наличие различных заимствований и делать определённые выводы о наличии этих заимствований.

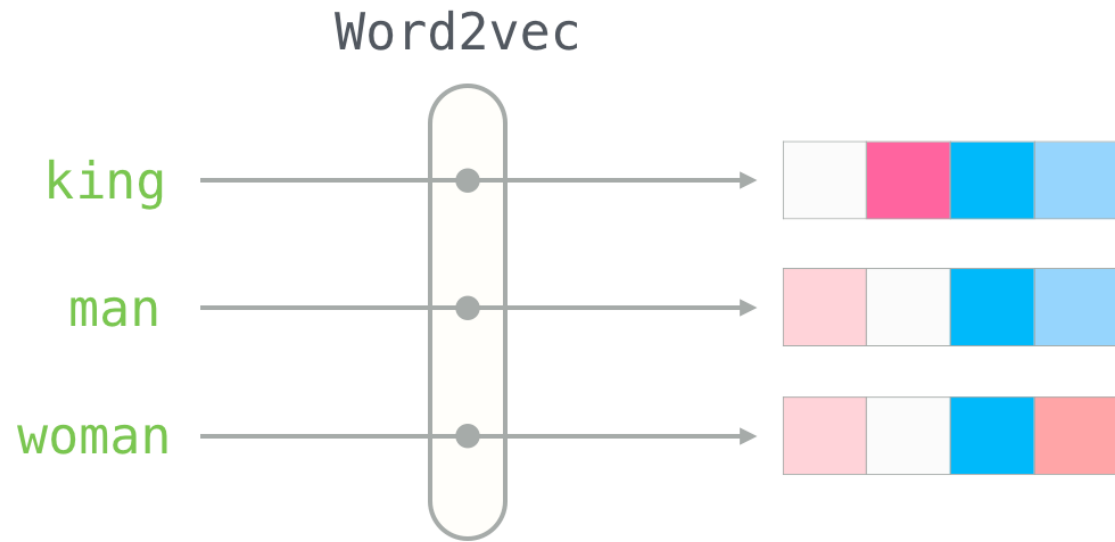
Задачи:

- сделать обзор аналогов разрабатываемой программы и имеющихся методов решения;
- определить требования к программе;
- спроектировать функциональную модель программы;
- реализовать и протестировать программу.

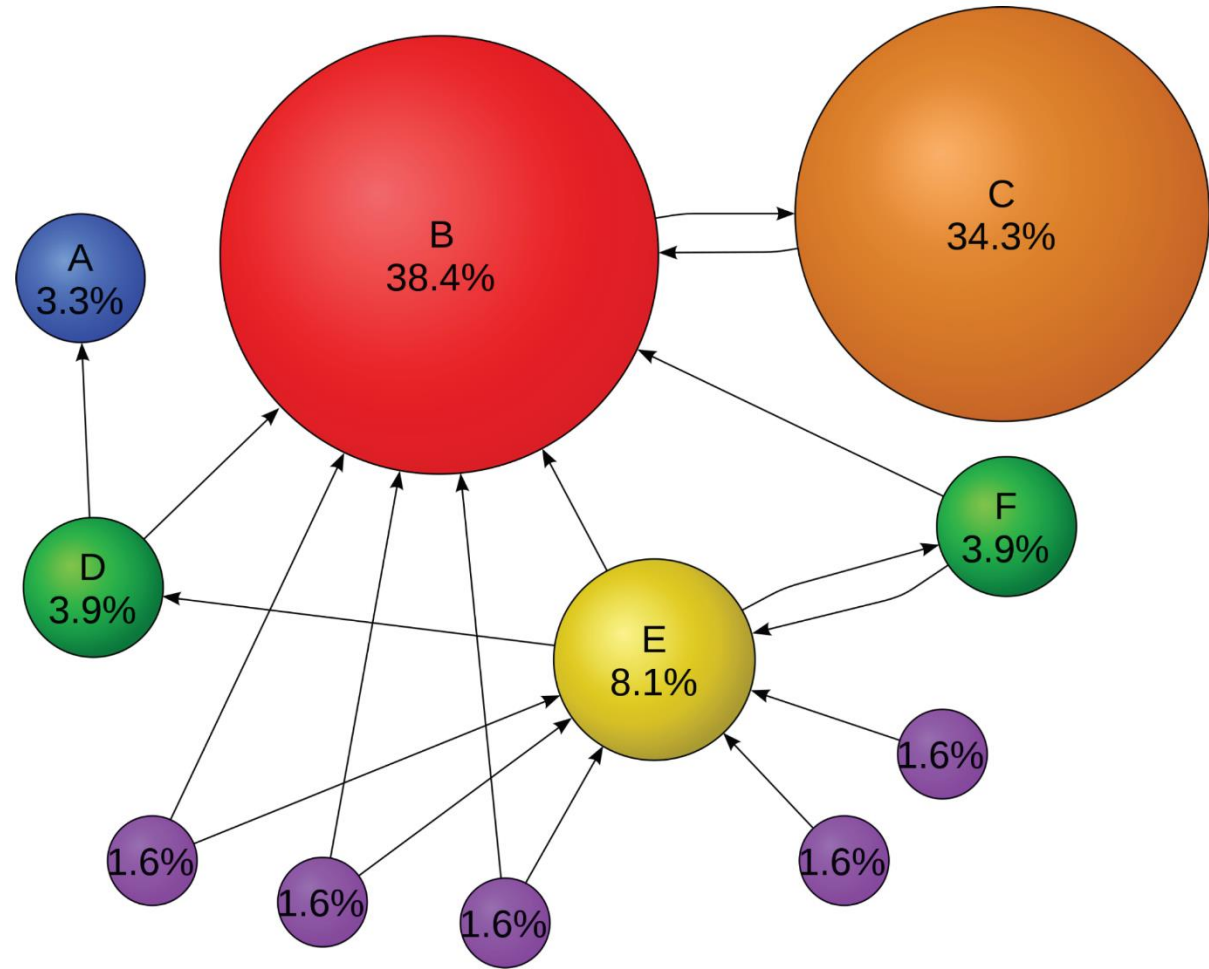
Список сервисов, использующихся для выявления плагиата в электронных таблицах:

1. Unichекk;
2. Plagiarism Checker X;
3. DupliChecker.

Анализ имеющихся решений

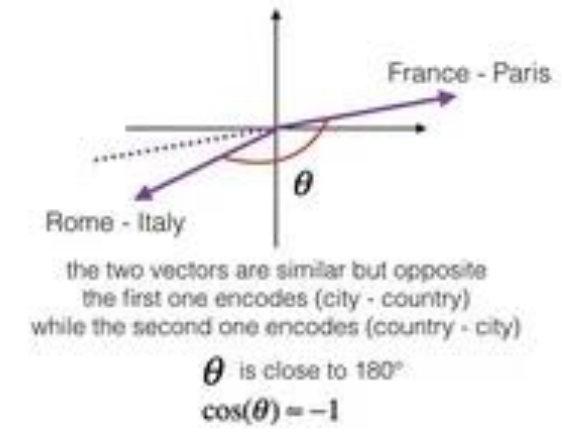
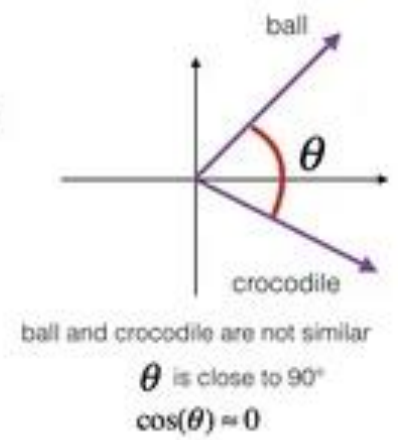
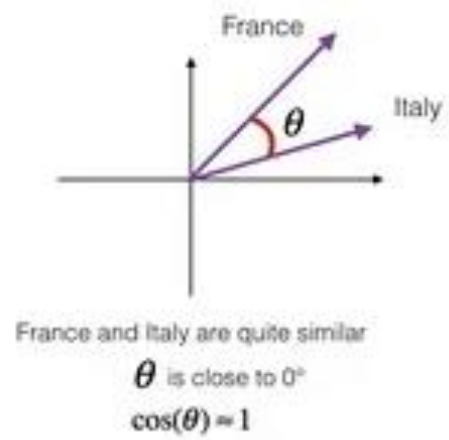
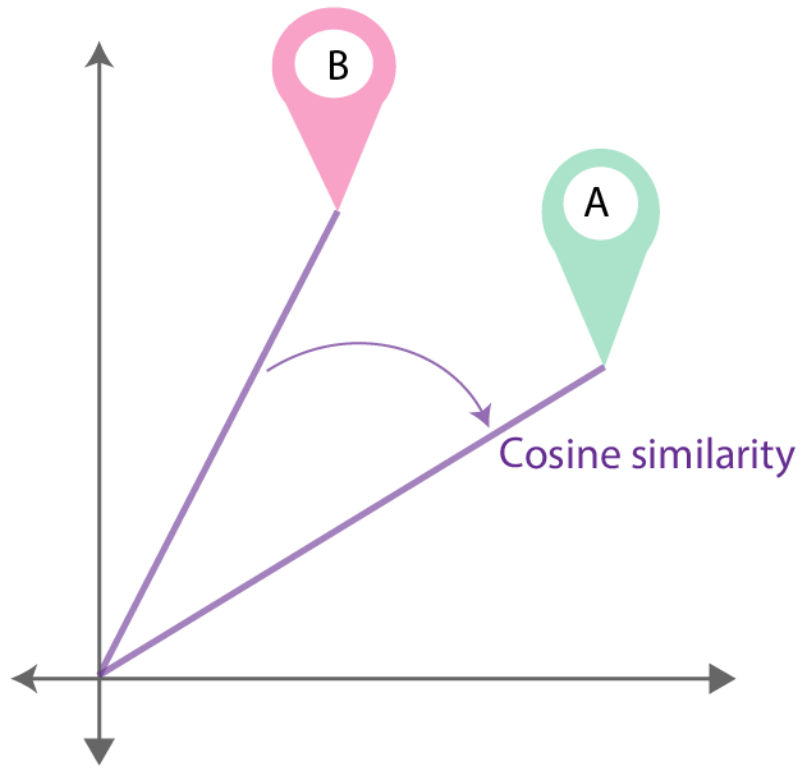


Метод векторного представления – Word2vec



Алгоритм ранжирования - PageRank

Косинусное сходство



Функциональные требования к программе

Программа должна соответствовать следующим требованиям:

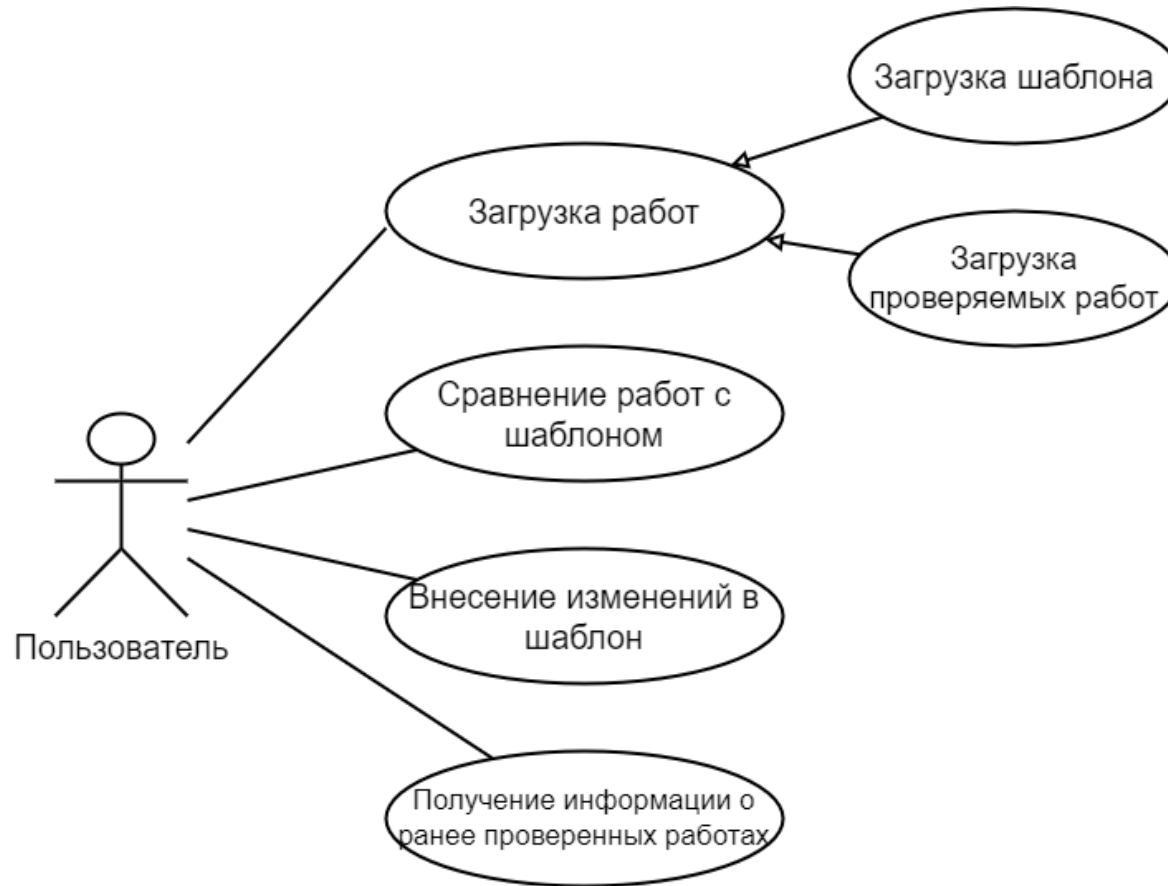
- 1) сопоставление проверяемой работы с образцом и формирование вывода о степени соответствия форматов и содержимого определённых ячеек;
- 2) показ информации обо всех обнаруженных несоответствиях;
- 3) возможность ручной проверки ячеек, информация в которых должна отмечаться в зависимости от исполнителя работы (содержит данные, которые будут уникальны для каждого исполнителя);
- 4) сохранение статистики проверенных работ (степень соответствия, обнаруженные ошибки);
- 5) наличие функции внесения незначительных изменений в исходный табличный формат;
- 6) выполнение анализа проверенных работ на наличие маркеров, указывающих на возможное заимствование.

Нефункциональные требования к программе ⁹

Нефункциональные требования к программе:

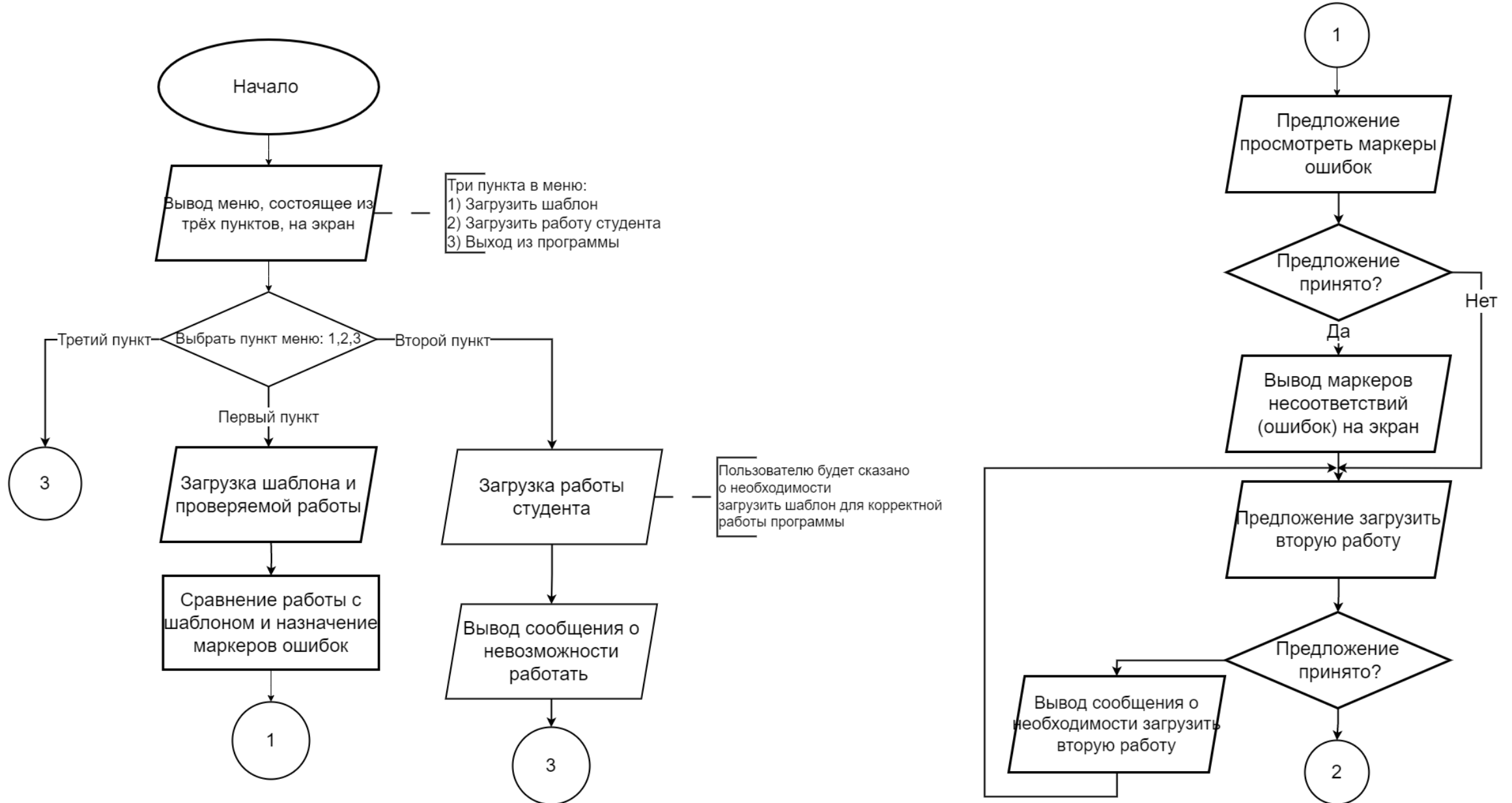
- 1) наличие кнопочного меню для удобной работы с программой;
- 2) программа должна работать на следующей ОС: Windows;
- 3) программа не должна требовать и собирать личные данные пользователя.

Проектирование функциональной модели программы

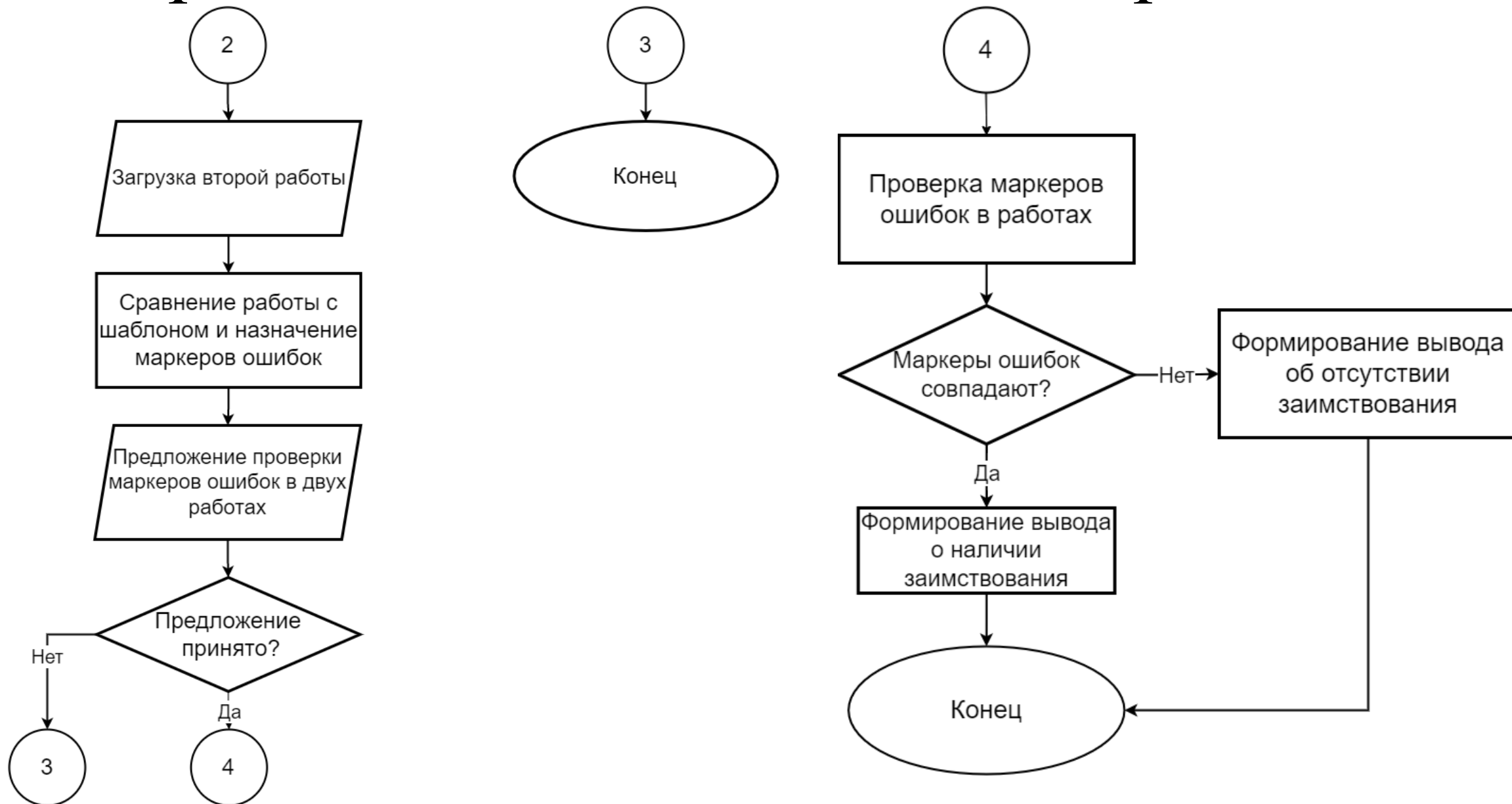


UML-диаграмма вариантов использования

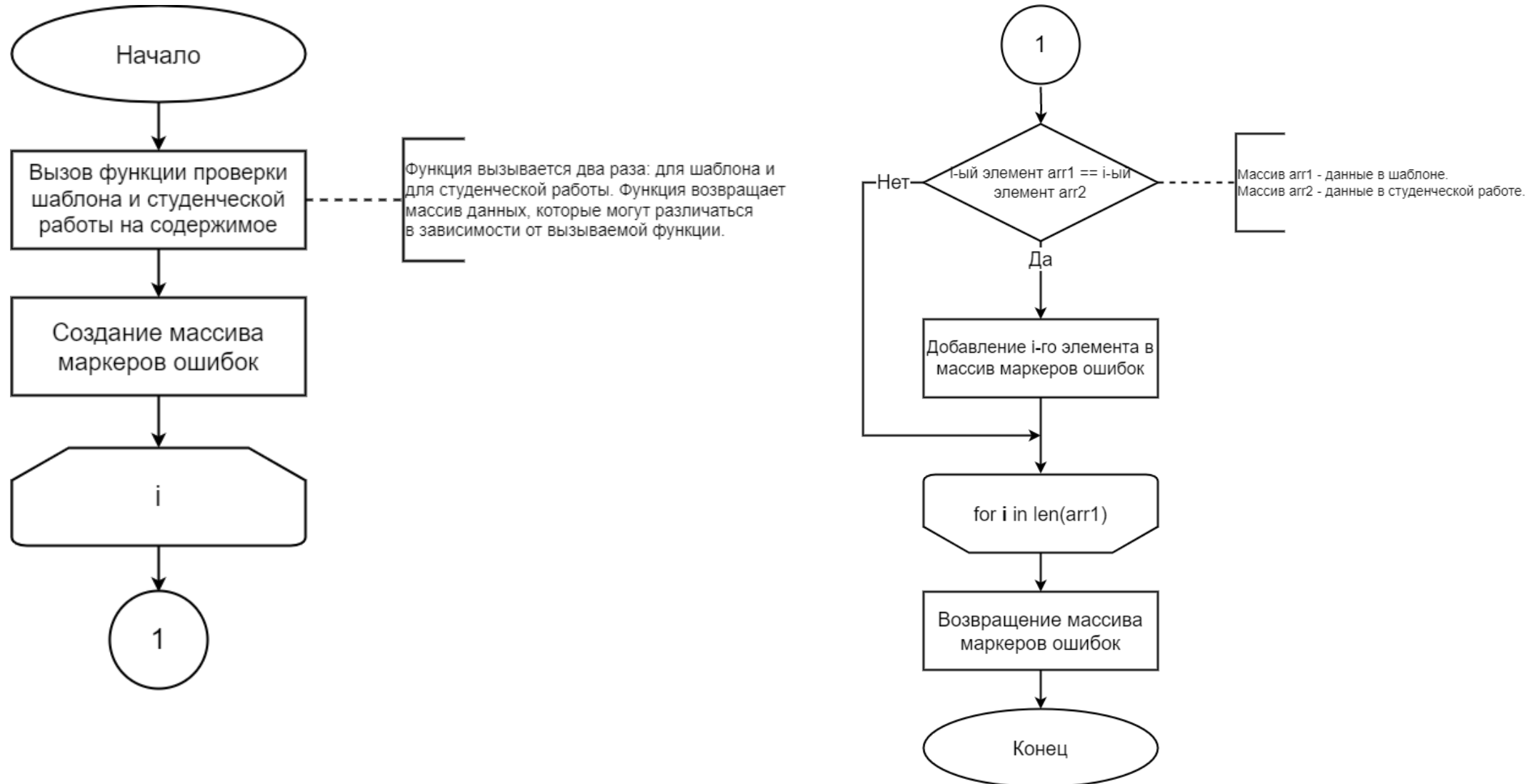
Блок-схема алгоритма работы программы



Продолжение блок-схемы алгоритма



Блок-схема алгоритма функции установления маркеров ошибок



Выбор средств реализации

Для разработки программы был выбран язык программирования Python в среде PyCharm Community Edition.

Для работы с электронными таблицами были использованы следующие библиотеки:

- pandas;
- openpyxl.

```
Menu:  
1 - Загрузить шаблон  
2 - Загрузить работу студента  
3 - Выход  
Выберите действие: |
```

Меню программы

```
Выберите действие: 1  
Укажите путь к Вашему шаблону  
  
Ваш путь: C:\Users\ADMIN\OneDrive\Рабочий стол\Учёба\ВКР\100 Шувалов 221 работа 1.xlsx  
Нажмите 2, если Вы хотите загрузить работу студента.  
Ваш выбор: |
```

Загрузка шаблона работы

Реализация. Формирование вывода о наличии/отсутствии заимствования

```
Сравнить ошибки двух работ? 1 - Да; 2 - Нет  
1  
Одинаковых ошибок нет.
```

Вывод об отсутствии одинаковых ошибок

```
Список одинаковых ошибок:  
9  
9  
9  
9  
9  
9  
9  
9  
9
```

Вывод о наличии одинаковых ошибок

Заключение

Целью работы являлась разработка программы для поиска заимствований по файлам электронных таблиц. Для достижения цели были решены следующие задачи:

- произведён обзор аналогов разрабатываемой программы и имеющихся методов решения;
- определены требования к программе;
- спроектирована функциональная модель программы;
- была реализована и протестирована программа.

Библиографический список

1 Unichesk [Электронный ресурс]. – URL: <https://dev.abcdef.wiki/wiki/Unichesk> (дата обращения 20.02.2023).

2 Программа для проверки на плагиат: Plagiarism Checker X [Электронный ресурс]. – URL: <https://kakdelateto.ru/programma-dlya-proverki-na-plagiat-plagiarism-checker-x/> (дата обращения: 20.02.2023).

3 Топ-5 самых надежных инструментов для проверки на плагиат [Электронный ресурс]. – URL: <https://smodin.io/ru/blog/best-plagiarism-checker-top-5-plagiarism-checker/> (дата обращения: 20.02.2023).

4 Word2vec [Электронный ресурс]. – URL: <https://habr.com/ru/articles/446530/><https://lala.lanbook.com/5-besplatnyh-analogov-antiplagiata-dlya-prepodavatelej> (дата обращения: 22.02.2023).

5 Word2Vec – Википедия [Электронный ресурс]. – URL: <https://ru.wikipedia.org/wiki/Word2vec> (дата обращения: 24.02.2023).

Спасибо за внимание