Учреждение образования

“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ ”

Кафедра интеллектуальный информационных технологий

Отчет по лабораторной работе №4

по курсу: “Естественно-языковой интерфейс интеллектуальных систем”

Выполнил студент группы 921702: Баранов А.Д.

Проверил: Крапивин Ю.Б.

Минск

2022

**Цель работы**: освоить на практике основные принципы машинного перевода документов.

**Вариант 4:** немецкий, английский.

**Требования к разрабатываемой системе:**

* на входе – естественно-языковой текст на входном языке, подлежащий процедуре машинного перевода;
* подсчитать количество слов во входном тексте, количество переведенных слов;
* на выходе: – перевод входного текста на выходной язык, упорядоченный по частоте встречаемости в тексте список слов и их переводов на выходной язык с грамматической информацией, построенное дерево синтаксического разбора предложений;
* интерфейс системы должен быть простым и доступным для пользователей любого уровня, содержать понятный набор инструментов.

**Реализация**

Система состоит из двух частей: серверной и клиентской. В серверной части находится код на языке Python, который переводит текст, подсчитывает слова и составляет синтаксическое дерево. В реализации использовались Google Translate API для перевода текста, библиотека spaCy для синтаксического разбора текста и библиотека nltk для подсчета частоты встречаемости слов.



Рис. 1. Запросы к серверу

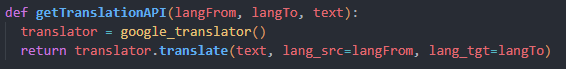


Рис. 2. Реализация перевода текста через Google Translate API

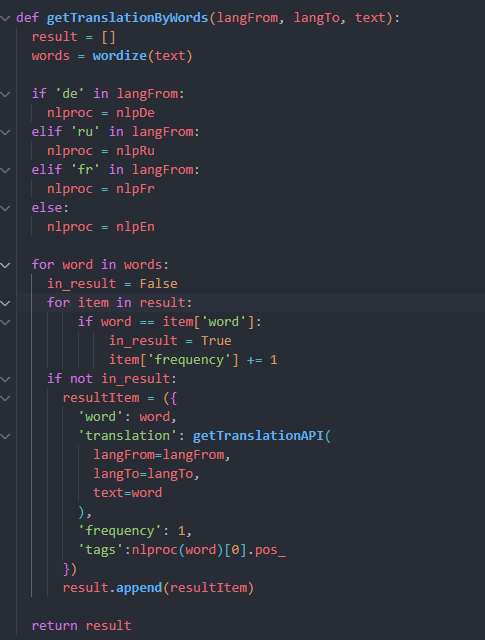


Рис. 3. Реализация подсчета слов в тексте

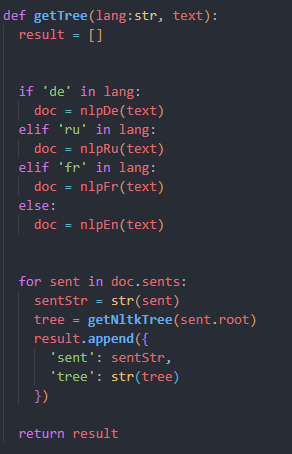


Рис. 4. Реализация составления синтаксического дерева

**Пользовательский интерфейс:**

Приложения написано на языке TypeScript с использованием фреймворка Angular.

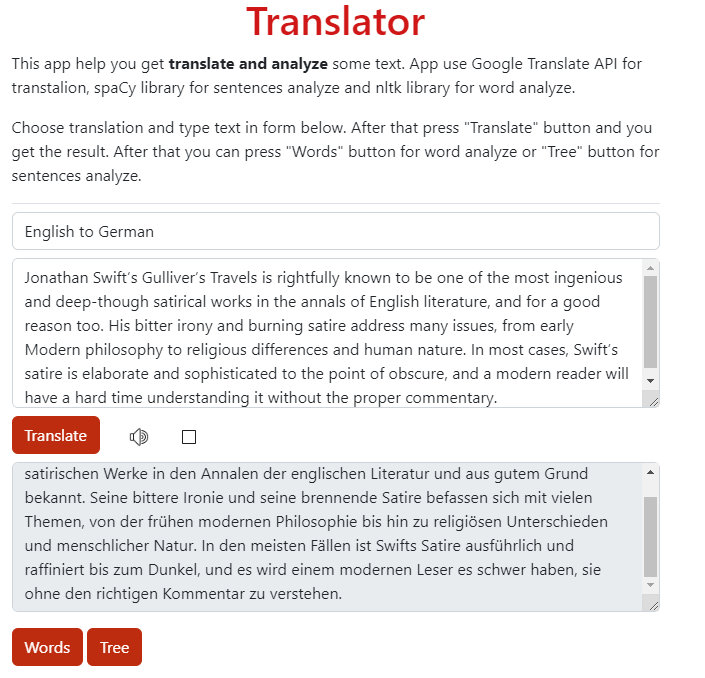


Рис. 5. Пользовательский интерфейс с результатом перевода

Пользователь вводит текст, который хочет перевести и нажимает кнопку “Translate”. После этого он получает результат в виде переведенного текста. Также пользователь может воспроизвести переведенный текст. Синтез речи реализован с использованием Web Speech API.

После перевода пользователь может воспользоваться кнопкой “Words”, чтобы отобразить таблицу со словами в тексте, или кнопку “Tree”, чтобы отобразить синтаксические деревья предложений переведенного текста



Рис. 8. Таблица частоты встречаемости слов

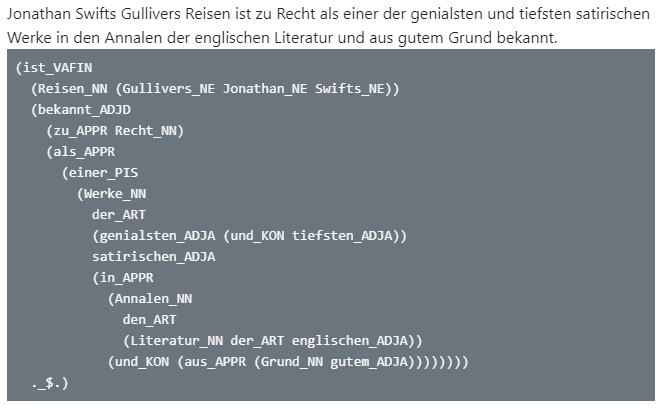


Рис. 9. Синтаксические деревья предложений

**Вывод:**

Разработанная система позволяет переводить текст и собирать некоторую информацию о нем: частоту встречаемости слов и синтаксические деревья. Система позволяет осуществлять перевод в два направления для 3-х языков: английский, французский и немецкий.