Учреждение образования

“БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ ”

Кафедра интеллектуальный информационных технологий

Отчет по лабораторной работе №4-5

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 создания систем анализа и синтеза речи.

**Вариант 6:** реализовать систему анализа, поддерживающую русский язык, и систему синтеза, поддерживающую русский язык.

**Требования к системе анализа:**

* задание списка операций, на которые система может реагировать;
* организация автоматической реакции системы на речевой сигнал с уведомлением пользователя о происходящем.

**Требования к системе синтеза:**

* ввод текста, или копирование через буфер, или поддержка указателя мыши в любом другом приложении;
* воспроизведение сгенерированного речевого сигнала для введенного текста.

**Реализация системы анализа**

Система состоит из двух частей: серверной и клиентской. Клиентская часть написана на языке С# и JavaScript на платформе ASP.NET. В серверной части находится код на языке Python, который обрабатывает текст.

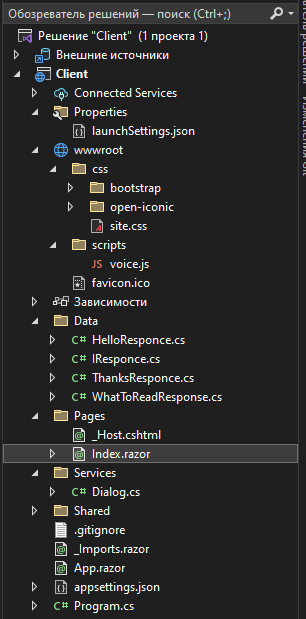


Рис. 1. Структура проекта

**Пользовательский интерфейс**

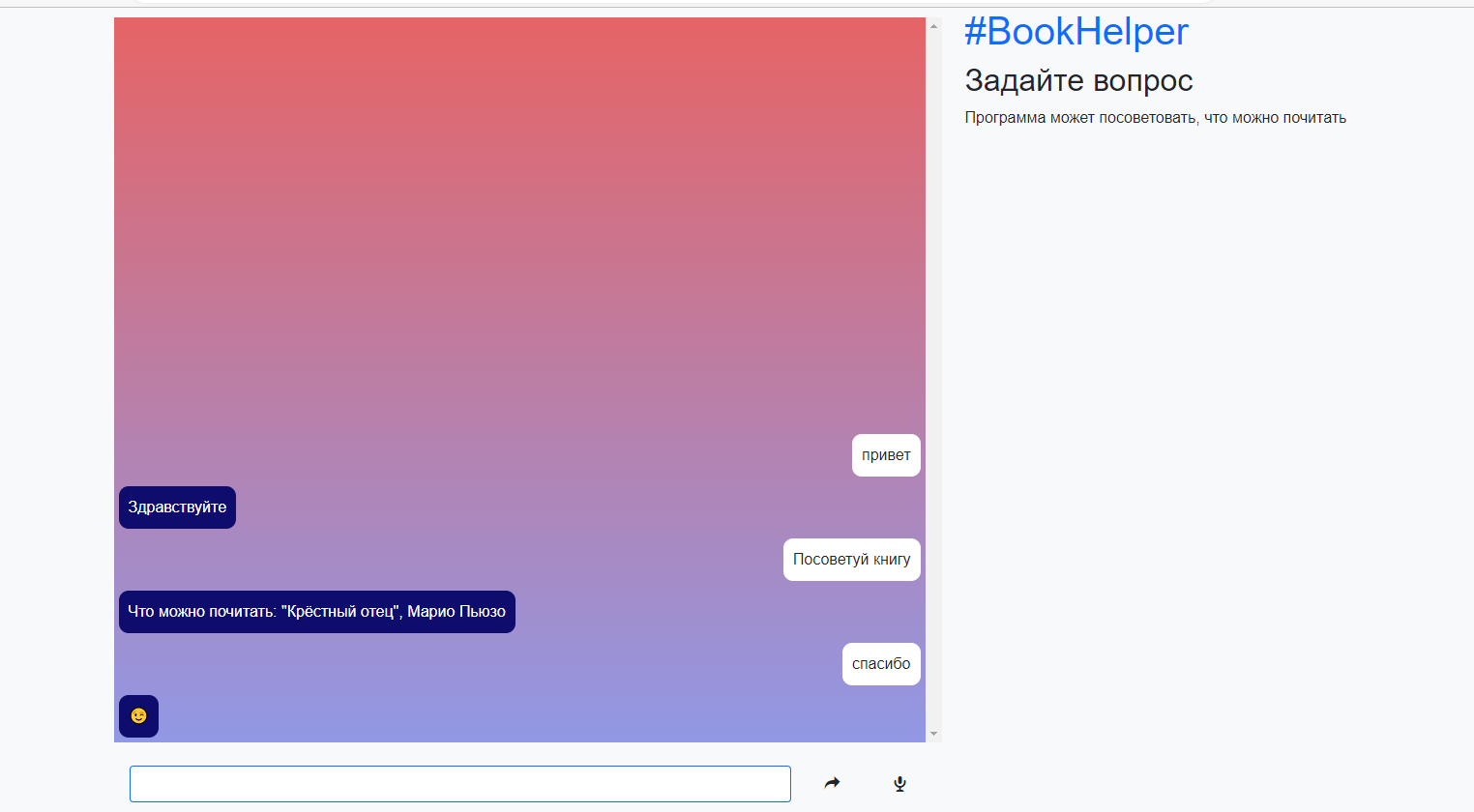


Рис. 2. Пользовательский интерфейс

Система анализирует голосовую команду или текстовую команду и дает соответствующий ответ.

**Системы обработки и поиска ответа**

Нужный ответ ищется по ключевым словам, которые генерируются в начале выполнения.



Рис. 3. модуль WhatToReadResponse

Каждый модуль реализует интерфейс IResponce. У каждого модуля есть список ключевых слов, которые получаются путем анализа нескольких главных слов (на Рисунке 5 этими словами являются “завтра” и “утро”).

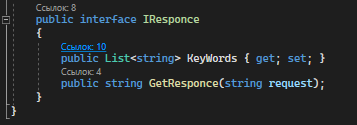


Рис. 4. IResponse

Диалоговая система отправляет каждому добавленному модулю текст пользователя. При первом соответствии с ключевым словом пользователю выдается один из заранее готовых ответов.

Диалоговая система взаимодействует с модулями через интерфейс, что дает возможность добавлять новые модули в систему. При это стоит учитывать в каком порядке модули добавляются в систему, для того чтобы получать наиболее ожидаемый ответ.

Для генерации ключевых слов из главный использовался модуль ruconceptnet.

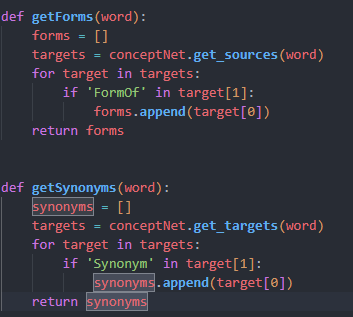


Рис. 5. Методы getForms и getSynonyms

Сервер находит все ключевые слова и отправляет приложению. После этого все слова сохраняются в приложении.

Текстовый ввод обрабатывается при заполнении текстового поля.



Рис. 6. Обработка текстового ввода

Голосовой ввод пользователя обрабатывается с помощью экспериментальной технологией Web Speech API языка JavaScript. Часть технологии, отвечающая за анализ голоса называется Speech Recognition.

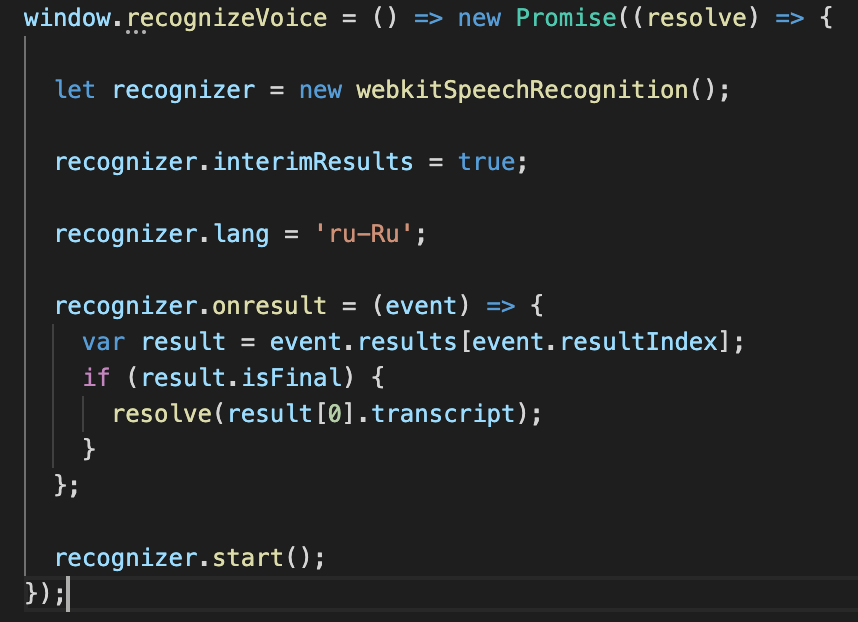


Рис. 8. Распознавание текста с помощью Web Speech API

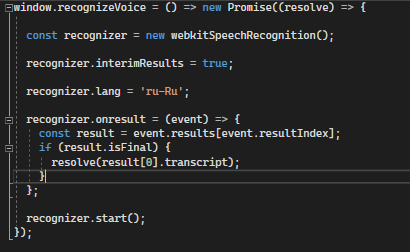


Рис. 7. Обработка голосового ввода

**Реализация системы синтеза**

Система представляет из себя веб-приложение, реализованное с помощью фреймворка Angular. Для синтеза речи используется экспериментальная технология Web Speech API.

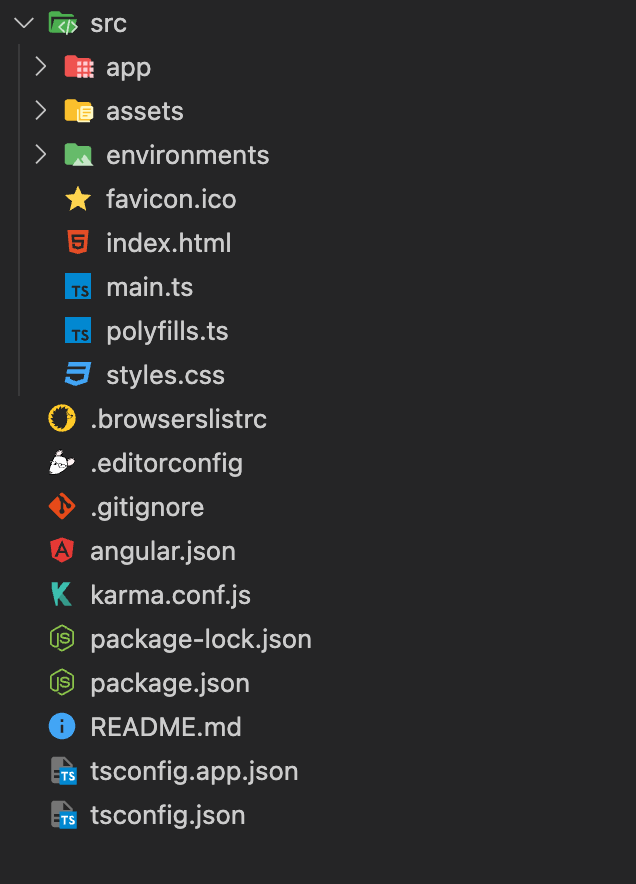


Рис. 9. Структура проекта

**Пользовательский интерфейс**

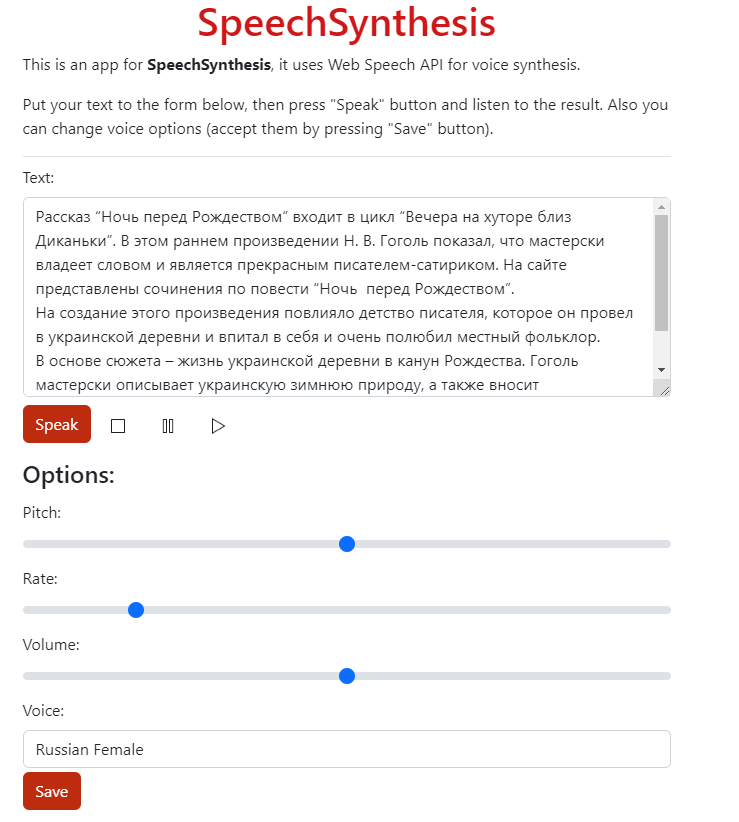


Рис. 10. Пользовательский интерфейс

Пользователь вводит текст и нажимает кнопку “Speak”. Система начинает воспроизведение текста. Пользователь может управлять воспроизведением: досрочно завершить воспроизведение, остановить и продолжить воспроизведение.

Пользователю доступны опции для настройки: высота голоса, темп, громкость и тип голоса (женский/мужской). По нажатию на кнопку “Save” текущее воспроизведение останавливается.

**Система синтеза голоса**

Часть Web Speech API, отвечающая за синтез речи, называется Speech Synthesis.

Для синтеза речи необходим объект SpeechSynthesisUtterance. Он содержит в себе высказывание, которое будет воспроизводится, а также настройки голоса. Сам синтез выполняет объект SpeechSynthesis.

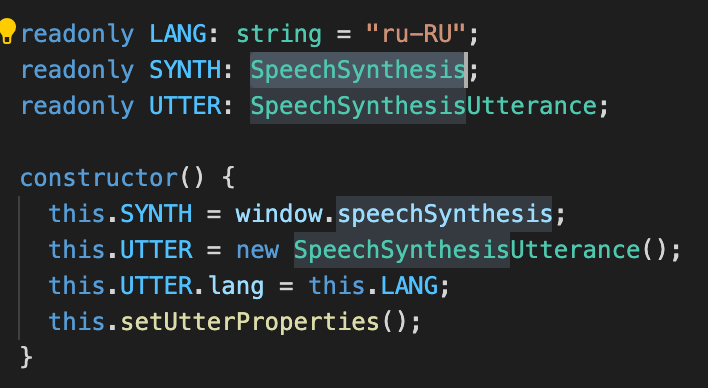


Рис. 11. Константные свойства и конструктор MainComponent

Основной метод класса MainComponent это метод TextToVoice, который запускается при нажатии на кнопку “Speak”.



Рис. 12. Метод TextToVoice

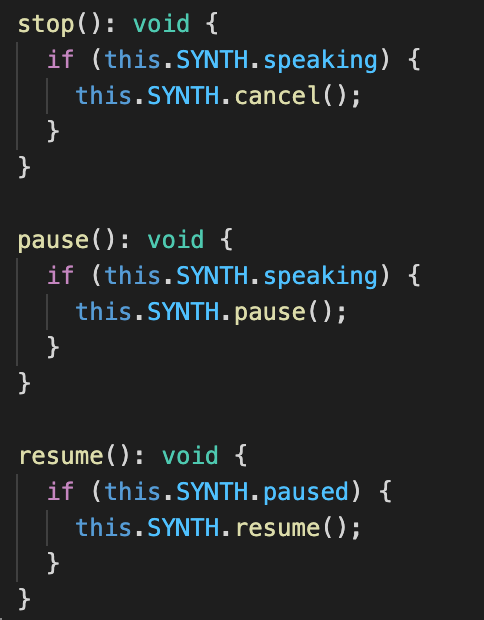


Рис. 13. Методы управления воспроизведением

**Вывод:**

В ходе работы были изучены и отработаны практические навыки применения инструментов для синтеза и анализа речи. Были созданы две системы на основе экспериментально технологии Web Speech API на разных фреймворках для создания веб-приложений. Т.к. обе системы разработаны на разных фреймворках и работают корректно, можно сказать, что технология Web Speech API может без каких-либо проблем использоваться в современных веб-приложениях.