

РАЗРАБОТКА IP-КОМПОНЕНТА СИНТЕЗА РЕЧИ

Литвинов А. А.

Кафедра интеллектуальных информационных технологий

Научный руководитель: Житко В.А., ассистент

e-mail: zhitko.vladimir@gmail.com

Аннотация — Анализируются существующие инструменты синтеза речи, описывается схема работы систем синтеза речи, а также рассматривается разработанный ip-компонент: схема взаимодействия модулей, используемые для реализации программные средства.

Ключевые слова: синтез речи, IP-компонент, интеллектуальная справочная система.

Синтез речи или TTS – это преобразование заранее неизвестной текстовой информации в речь [1]. Он предоставляет еще один канал передачи данных от электронных устройств к человеку наравне со зрительным восприятием. Технология синтеза устной речи находит широкое применение у людей, имеющих проблемы со зрением. Для всех остальных она создает новое измерение удобства пользования техникой и значительно снижает нагрузку на зрение, на нервную систему, позволяет задействовать слуховую память. Данная технология может быть использована для чтения информации из электронных ресурсов, в навигационных системах и автоматических переводчиках. Любой текст состоит из слов, разделенных пробелами и знаками препинания. Произнесение слов зависит от их расположения в предложении, а интонация фразы – от знаков препинания. Наконец, произнесение зависит и от смысла слова. Соответственно, для того, чтобы синтезированная речь звучала естественно, необходимо решить целый комплекс задач, связанных как с обеспечением естественности голоса на уровне плавности звучания и интонации, так и с правильной расстановкой ударений, расшифровкой сокращений, чисел, аббревиатур и специальных знаков с учетом особенностей грамматики языка. На рынке программного обеспечения в области синтеза речи существует достаточное количество решений, поэтому для достижения цели, которой является разработка ip-компонента для синтеза речи, в качестве основы был выбран уже готовый программный продукт. Для достижения вышеописанной цели были поставлены следующие задачи:

- проанализировать существующие системы синтеза речи;
- выделить наиболее качественное решение;
- разработать ip-компонент на основе выбранного программного продукта.

Среди существующих систем синтеза речи можно выделить следующие: Festival, MBROLA, Flite, FreeTTS, VaBel Technologies, NeoSpeech. Самой популярной и качественной является система Festival,

которая является разработкой Исследовательского Центра Речевых Технологий (CSTR) университета Эдинбурга. Из достоинств этой системы можно выделить следующие: поддержка большого количества языков, кроссплатформенность, качество произношения. Система представляет собой набор исполняемых файлов и библиотек языков. На ее вход подается текст в электронном виде, на выходе получается звуковой файл в формате WAV.

Для интеграции системы Festival в систему OSTIS был использован язык программирования Python. Доступ каждого IP-компонента к интеллектуальной справочной системе (ИСС) происходит через API (Интерфейс программирования приложений), который предоставляет доступ к ядру ИСС и дает возможность пользоваться ее ресурсами. Таким образом, взаимодействие системы синтеза речи и ИСС происходит через прослойку, которая и представляет собой IP-компонент. Диаграмма последовательностей представлена на рисунке 1.

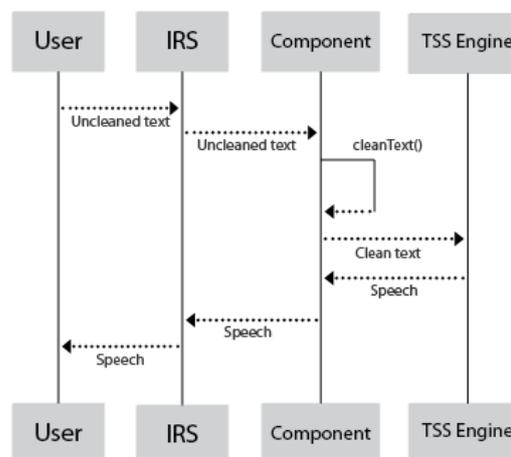


Рис. 1. Диаграмма последовательностей

«Неочищенный текст» подается на вход ИСС, справочная система передает этот текст на вход IP-компонента. Компонент очищает текст от повторяющихся слов, знаков препинания и т.д. Далее текст подается на вход системы синтеза речи, которая возвращает сгенерированную по тексту речь. Она переправляется обратно по цепочке и на выходе пользователь слышит необходимую фразу. Данный подход позволяет с легкостью заменить составляющие этой схемы. Например, можно существующую систему синтеза речи заменить другой, более качественной, если такая появится. Такой гибкости добиваются благодаря максимальной обособленности функционально-законченных модулей системы.

В качестве направления дальнейшего развития можно рассматривать расширение списка поддерживаемых языков.