

УДК 004.8
ВОЗМОЖНОСТИ АВТОМАТИЗАЦИИ ОБРАБОТКИ ДОКУМЕНТОВ
В MS WORD

Д. В. ЗДАНОВИЧ, Н. В. ВЫГОВСКАЯ, В. М. ПРУДНИКОВ
Государственное учреждение высшего профессионального образования
«БЕЛАРУССКО-РОССИЙСКИЙ УНИВЕРСИТЕТ»
Могилев, Беларусь

Текстовый процессор MS Word из пакета MS Office имеет множество функций для создания, просмотра и редактирования документов. Однако опытными пользователями востребованы комплексные инструменты, позволяющие автоматизировать обработку документа, используя возможности MS Word.

Известные пакеты таких программ предоставляют широкие возможности по обработке документов, но используя объектную модель документа и язык VBA, могут работать только с определенными версиями MS Word (например “Перестройка 2002” с версиями MS Word 2000/2002).

Разработано приложение в виде самостоятельного модуля на языке Visual Basic.NET. В приложении используется технология взаимодействия с ActiveX-объектами через объектную модель DOM MS Word. Приложение по команде пользователя получает ссылку на ActiveX-объект «Word.Application», что позволяет глубоко обрабатывать документы, как и при использовании VBA.

Реализация такого подхода позволила получить следующие преимущества:

- работоспособность на различных версиях MS Word;
- возможность работы одновременно с несколькими документами (приложение не имеет привязки к конкретному документу);
- быстрая автоматизированная обработка всего документа или его части (нет необходимости настраивать параметры для каждого участка текста);
- использование сторонних средств и технологий, например, технологии озвучивания текста (Text-to-Speech);
- портабельность.

В настоящее время программа обеспечивает следующие функции:

- мультивозврат – запоминание состояния документа и, при необходимости, выполнение возврата к нему;
- перестройка текста – исправление типичных ошибок набора (удаление лишних символов конца абзаца (¶), упорядочивание пробелов в скобках и сокращениях, короткие строки и т.д.);
- обработка ошибочных и лишних символов, полученных после использования OCR-программ;
- просмотр шрифтов, оптимизация размеров абзаца;

- работа с определением языка слов (поиск слов, состоящих из смеси русских и английских букв, их анализ и замена английских букв в русских словах на русские и наоборот);
- распечатка документа в виде тетради;
- чтение текста – используется встроенная в MS Windows технология Text-to-speech engine.

Интерфейс приложения реализован в виде панели с набором виртуальных кнопок для вызова соответствующих инструментов (рис. 1).



Рис. 1. Интерфейс приложения

Апробация разработанной программы показала её практическую востребованность. Разработка и дополнение инструментов продолжается.