

База знаний интеллектуальной справочной системы по русскому языку

УО «Белорусский государственный университет информатики и радиоэлектроники»
г. Минск, Республика Беларусь

Мысовских Д.С., Аврамова А.И., Пылинский В.И., Пылинский В.И.

Житко В. А. – ассистент каф. ИИТ

Сегодня особый интерес в области высоких технологий стали привлекать программные системы призванные заниматься обработкой и обслуживанием естественно-языковой информации. Важными в этой категории являются проблемы анализа и синтеза речевых конструкций, обработки текстов, обучения персонала, а также детских ресурсов в данной области. Для решения этих сложных задач необходимо иметь прочный базис, который заложит интеллектуальная справочная система по русскому языку.

Успех в решении прикладных задач компьютерной лингвистики зависит, прежде всего, от полноты и точности представления в памяти ЭВМ декларативных средств. Таким образом, проблема проектирования лингвистических баз знаний наиболее полно описывающих естественный язык стоит наиболее остро. На сегодняшний день необходимый уровень решения этих задач пока еще не достигнут, хотя работы в области компьютерной лингвистики ведутся во всех развитых странах мира.

Лингвистическая база знаний интеллектуальной системы по русскому включает в себя полное описание естественного языка, что может включать в себя: фонетические, графические, семантические, лексические, грамматические (морфологические и синтаксические) правила, а также различные лингвистические словари.

Лингвистические правила, а также словари, записываются на едином специализированном языке, разработанном для описания различных лингвистических структур. Использование семантических сетей для «хранения» словарей естественных языков позволяет:

1. избавиться от избыточности данных;
2. объединять словари в одну базу знаний.

Такая база знания представляет собой отдельный компонент, который будет востребован как в различных интеллектуальных справочных системах в качестве предметной базы знаний, так и в системах по обработке естественно-языковых текстов.

В базе знаний по русскому языку можно выделить следующие направления:

- лексика и фразеология - описывает словарный и фразеологический состав русского языка и закономерности его развития;
- фонетика - описывает звуковой состав современного литературного русского языка и основные звуковые процессы, протекающие в языке;
- орфоэпия - описывает нормы современного русского литературного произношения;
- графика - описывает состав русского алфавита, соотношения между буквами и звуками. В нее так же включается орфография, описывающая основные принципы русского написания - морфологические, а так же фонетические и традиционные;
- морфемика и словообразование - описывает морфологический состав слова и основные типы образования новых слов.
- грамматика – включает морфология (описывает грамматические категории и грамматические формы слова) и синтаксис (описывает основные синтаксические единицы - словосочетания и предложения, виды синтаксической связи, типы предложений и их структуру);
- пунктуация - описывает совокупность правил расстановки знаков препинания;
- текст, признаки, характеристики и т.д.

Основной задачей интеллектуальных систем является предоставление информации пользователю по его запросу. Возможность и правильность ответа системы, в первую очередь, будет зависеть от качества спроектированной базы знаний.

Список источников:

- Д.Э. Розенталь, И.Б. Голуб Русский язык / Учебное пособие 2-е изд., перераб. и доп. - М.: Рольф, 2001. - 382 с.
- Л.А. Шевченко, Н.М. Пипченко Пособие по русскому языку и литературе для поступающих в вузы - М.: БГУ, 1975. - 320с.
- Н.И. Гурский, Л.М. Филипович, А.М. Бордович Русский язык, часть 1, фонетика и морфология - М.: Народная асвета, 1970. - 317с.
- Л.С. Мормыш, Т.А. Павлюченко. Пособие-репетитор, Русский язык -Минск 2007
- Р.Г. Четчет Русский язык / Учебное пособие 2-е изд., перераб. и доп. - М.: Университэцкае, 1999. - 301 с.
- Горбатов В.А. Фундаментальные основы дискретной математики. Информационная математика. - М.: Наука, Физматлит, 2000.- 544 с.
- Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженеров / Учебное пособие 3-е изд., перераб. и доп. - М.: Лань, 2004 г. - 400с.
- Новиков Ф.А. Дискретная математика для программистов / Учебное пособие 2-е изд., перераб. и доп. - СПб.: Питер, 2003 г. - 364 с.
- Математическая логика: Учеб. пособие / Л.А.Латонин, Ю.А.Макаренко, В.В.Николаева, А.А.Столяр. Под общ.ред. А.А.Столяра. - Мн.: Выш. школа, 1991. - 269с.
- Open Semantic Technology for Intelligent Systems[Электронный ресурс] / Ostis Минск, 2010 <http://ostis.net>