

# Интеллектуальная справочная система по русскому языку

Нечипуренко Д.И.; Русецкий К.В.; Скиба Е.В.; Снигурова И.В.

Каф. ИИТ, ФИТиУ

Белорусский государственный университет информатики и радиоэлектроники

Минск, Беларусь

e-mail: tin-moon@mail.ru

**Аннотация** – Основной целью данной работы является разработка интеллектуальной вопросно-ответной системы по теории русского языка, включающую в себя машину обработки знаний для надления интеллектуальной справочной системы по русскому языку возможностью генерировать знания, в явном виде не содержащихся в базе знаний.

**Ключевые слова:** интеллектуальные системы, русский язык, лингвистика, базы знаний

## I. ВВЕДЕНИЕ

В настоящее время особый интерес в области информационных технологий стали представлять программные продукты и средства 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.

Основной задачей интеллектуальных систем является предоставление информации пользователю по его запросу. Основными средствами, при этом, являются операции навигации и поиска по семантическим сетям. Главной особенностью таких средств является то, что они могут использовать различные подходы для поиска нужной информации (подход с поиском по шаблону или интеллектуальным поиском и др.), но при этом должны быть интегрированы в одну систему и использовать один источник данных.

## II. ОБЩАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОЙ СПРАВОЧНОЙ СИСТЕМЫ

Интеллектуальная справочная система состоит из трех компонентов:

1. База знаний [1], представленная в виде семантической сети.
2. Машина обработки знаний, представленная агентами обработки семантической сети.
3. Пользовательский интерфейс использует для общения с пользователем подмножество SCg языков [2].

В интеллектуальных системах [1] информация представляется в виде семантической сети, что позволяет оперировать не только фактографической информацией, но и осуществлять навигацию по установленным отношениям [3,4,5] в рамках предметной области прикладной вопросно-ответной системы. Пример представления фрагмента лингвистической базы знаний в виде семантической сети изображен на рисунке 1. Исходной формой представления знаний является ограниченный естественный язык, пример статьи на котором представлен на рисунке 2. Лингвистические знания о естественном языке, записанные в лингвистической базе знаний, используются для анализа естественно-языкового текста, а также для его синтеза.

Лингвистическая база знаний также используется для решения и генерации заданий для пользователя и его контроля и проверки. Эта часть базы знаний выделяется в качестве предметной базы знаний по русскому языку. Отметим также важность интеллектуальных вопросно-ответных, т.к. они составляют основу интеллектуальных систем.

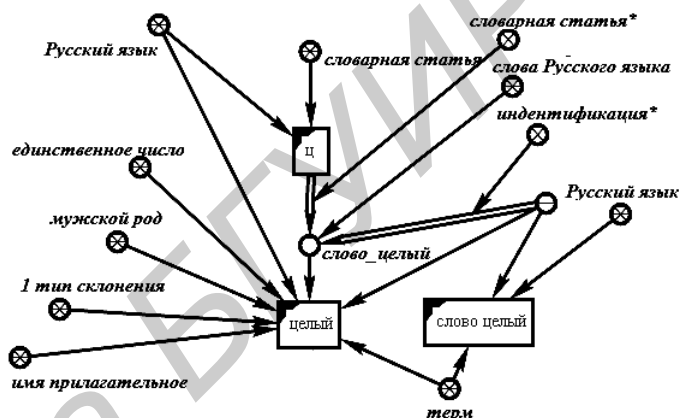


Рис. 3. Запись словарной статьи на языке SCg

### имя существительное

- = существительное
- = множество всех существительных
- = множество всех имен существительных
- ∈ Лингвистика русского языка
- ∈ множество
- ⊂ часть речи
- ⊂ самостоятельная часть речи
- ⊃ собирательное существительное
- Р а з б и е н и е (по собственности-нарицательности):
  - собственное существительное
  - нарицательное существительное
- Р а з б и е н и е (по одушевленности):
  - одушевленное существительное
  - неодушевленное существительное
- Р а з б и е н и е (по образованию):
  - отглагольное существительное
  - отпричастное существительное
- П р и м е р ы :
  - солнце
  - подснежник

Рис. 4. Описание понятия предметной области на псевдоестественном языке

Машина обработки знаний представлена информационно-поисковыми операциями и операциями генерации новых знаний (вывод по аналогии, прямой вывод, обратный вывод [6]). Информационно-поисковые операции используются для навигации пользователя по семантической сети, поиска ответов на заданные вопросы и пр. Например [7,8,9,10,11]:

1. Поиск определения понятия

2. Поиск примеров для понятия
3. Склонение имени существительного
4. Склонение имени прилагательного
5. Образование краткой формы прилагательного
6. Вывод семантической окрестности узла

Операции генерации новых знаний используются для получения новых знаний на основе существующих. К таким операциям могут относиться: генерация задания пользователю по определенной тематике, анализ выполненного пользователем задания (требуется провести анализ выполненного задания с целью проверки корректности). Пример результата разбора вопросительного предложения в виде семантической сети представлен на рисунке 3.

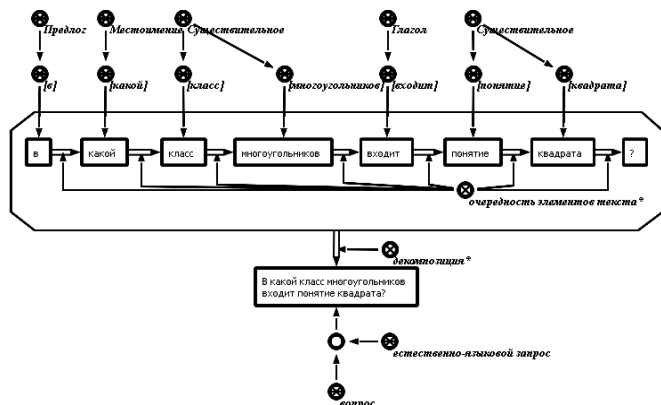


Рис. 5. Представление разобранного предложения в виде SCg-конструкции

Пользовательский интерфейс интеллектуальной справочной системы в рамках технологии OSTIS [12] представляет собой специализированную интеллектуальную систему, осуществляющую диалог с пользователем. Для интеллектуальной справочной системы по русскому языку пользовательский интерфейс может сочетать в себе элементы графического и естественно-языкового интерфейсов[2]. Графический – потому, что элементы интерфейса исполнены в виде графических изображений, которые отображают их назначение и свойства, что облегчает понимание и освоение программ неподготовленными пользователями. А естественно-языковой – потому, что он обладает следующими преимуществами:

- Минимальная предварительная подготовка пользователя, т.к. естественный язык является наиболее привычным и удобным средством коммуникации.

[14]

- Простота задания запросов на естественном языке.
- Большая скорость создания произвольного запроса, т.е. пользователь сразу может сформулировать корректное естественно-языковое представление запроса, поскольку такое представление является самым естественным для человека.

Пользовательский интерфейс интеллектуальной системы складывается из множества[13] взаимосвязанных и унифицированным образом взаимодействующих ip-компонентов пользовательских интерфейсов [12]. Унификация взаимодействия компонентов обеспечивает легкое расширение интерфейса новыми формами представления информации, наиболее адекватными каждой конкретной предметной области

[1] Математическая логика: Учеб. пособие / Л.А.Латонин, Ю.А.Макаренков, В.В.Николаева, А.А.Столяр. Под общ.ред. А.А.Столяра. - Мн.: Выш. школа, 1991. - 269с.

[2] Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский – СПб: Питер, 2000. – 384с.

[3] Горбатов В.А. Фундаментальные основы дискретной математики. Информационная математика. - М.: Наука, Физматлит, 2000.- 544 с.

[4] Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженеров / Учебное пособие 3-е изд., перераб. и доп. - М.: Лань, 2004 г. - 400с.

[5] Новиков Ф.А. Дискретная математика для программистов / Учебное пособие 2-е изд., перераб. и доп. - СПб.: Питер, 2003 г. - 364 с.

[6] Математическая логика: Учеб. пособие / Л.А.Латонин, Ю.А.Макаренков, В.В.Николаева, А.А.Столяр. Под общ.ред. А.А.Столяра. - Мн.: Выш. школа, 1991. - 269с.

[7] Д.Э. Розенталь, И.Б. Голуб Русский язык / Учебное пособие 2-е изд., перераб. и доп. - М.: Рольф, 2001. - 382 с.

[8] Л.А. Шевченко, Н.М. Пипченко Пособие по русскому языку и литературе для поступающих в вузы - М.: БГУ, 1975. - 320с.

[9] Н.И. Гурский, Л.М. Филипович, А.М. Бордович Русский язык, часть 1, фонетика и морфология - М.: Народная асвета, 1970. - 317с.

[10] Л.С. Мормыш, Т.А. Павлюченко. Пособие-репетитор, Русский язык -Минск 2007

[11] Р.Г. Чечет Русский язык / Учебное пособие 2-е изд., перераб. и доп. - М.: Университэцае, 1999. - 301 с.

[12] Open Semantic Technology for Intelligent Systems[Электронный ресурс] / Ostis Минск, 2010 <http://ostis.net>

[13] Харари Ф. Теория графов. Пер. с англ. 3-е изд. - М.: КомКнига, 2006. - 296с.