

ИНТЕЛЛЕКТУАЛЬНАЯ СПРАВОЧНАЯ СИСТЕМА ПО РУССКОМУ ЯЗЫКУ

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

Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь, tin-moon@mail.ru

Abstract. The primary objective of the present work is to develop an intelligent question answering system on the theory of the Russian language, which includes a knowledge processing machine that gives the intelligent reference system on the Russian language an ability to generate pieces of knowledge which are not explicitly present in the current state of the knowledge base.

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 (ИСС) 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.

ИСС состоит из трех компонентов:

- База знаний (БЗ) [1], представленная в виде семантической сети.
- Машина обработки знаний (МОЗ), представленная агентами обработки семантической сети.

- Пользовательский интерфейс (ПИ) 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 БЗ 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 БЗ, используются для анализа естественно-языкового текста, а также для его синтеза. Лингвистическая БЗ также используется для решения и генерации заданий для пользователя и его контроля и проверки. Эта часть базы знаний выделяется в отдельный компонент и может использоваться в качестве предметной базы знаний по русскому языку. Отметим также важность интеллектуальных вопросно-ответных систем, т.к. они составляют основу интеллектуальных систем.

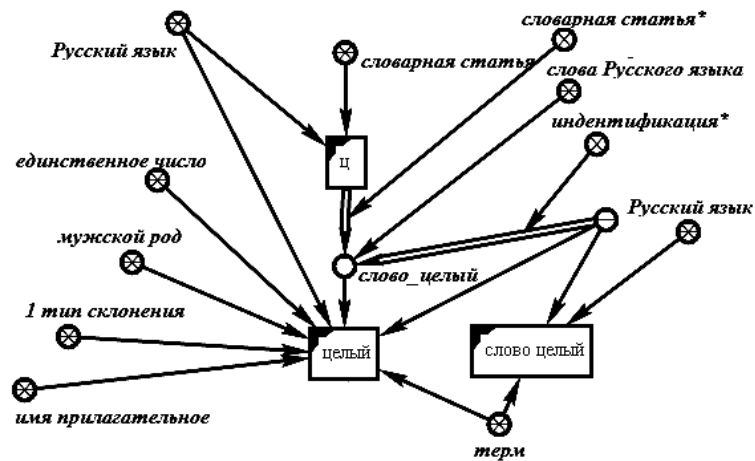


Рисунок 1 - Запись словарной статьи на языке SCg

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
- Примеры:
 - *солнце*
 - *подснежник*

Рисунок 2 - Описание понятия предметной области на псевдоестественном языке

МОЗ представлена информационно-поисковыми операциями и операциями генерации новых знаний. Информационно-поисковые операции используются для навигации пользователя по семантической сети, поиска ответов на заданные вопросы и пр. Например [6,7,8]: поиск определения понятия, поиск примеров для понятия, 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. Пример результата разбора вопросительного предложения в виде семантической сети представлен на рисунке 3.

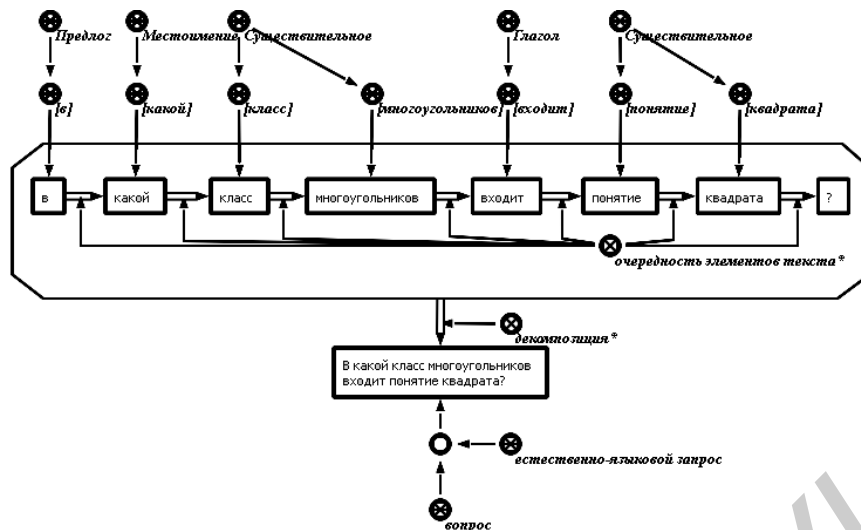


Рисунок 3 - Представление разобранного предложения в виде SCg-конструкции

ПИ ИСС в рамках технологии OSTIS [9] представляет собой специализированную интеллектуальную систему, осуществляющую диалог с пользователем. Для ИСС по русскому языку ПИ 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.
- Простота задания запросов на естественном языке.

ПИ ИСС складывается из множества[10] взаимосвязанных и унифицированным образом взаимодействующих ip-компонентов[9]. Унификация взаимодействия компонентов обеспечивает легкое расширение интерфейса новыми формами представления информации.

Литература

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. Д.Э. Розенталь, И.Б. Голуб Русский язык / Учебное пособие 2-е изд., перераб. и доп. - М.: Рольф, 2001. - 382 с.
7. Л.С. Мормыш, Т.А. Павлюченко. Пособие-репетитор, Русский язык -Минск 2007
8. Р.Г. Чечет Русский язык / Учебное пособие 2-е изд., перераб. и доп. - М.: Университэцкае, 1999. - 301 с.
9. Open Semantic Technology for Intelligent Systems[Электронный ресурс] / Ostis Минск, 2010 <http://ostis.net>
10. Харари Ф. Теория графов. Пер. с англ. 3-е изд. - М.: КомКнига, 2006. - 296с.