

## АЛГОРИТМ ПОСТРОЕНИЯ СЕМАНТИЧЕСКОЙ СЕТИ С ОБРАБОТКОЙ МЕСТОИМЕНИЙ ТЕКСТА

*Потараев В.В.*

*Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь*

*Серебряная Л.В. – к.т.н., доцент*

В современных информационных системах обрабатываются большие объёмы информации. При автоматизации работы с информационной системой может быть использован семантический анализ текстовой информации. Модель семантической сети является одним из инструментов семантического анализа. Местоимение – это часть речи, которая встречается в тексте довольно часто. Рассмотрим метод построения семантической сети с учётом местоимений текста.

В информационных системах обрабатываются довольно большие объёмы текстовой информации. Поэтому огромное значение имеет автоматизированная обработка информации, а также её быстрота и правильность. Применение семантического анализа и баз знаний является одним из инструментов автоматизированной обработки информации.

Местоимение является частью речи, которая позволяет упростить высказывания и используется довольно часто.

Целью данной работы является разработка метода автоматизированного построения семантической сети, который учитывает высказывания текста, сформулированные с помощью местоимений.

База знаний – это компонент экспертной системы, предназначенный для хранения долгосрочных данных, описывающих определенную предметную область, и правил, описывающих целесообразные преобразования данных этой области [1].

Выделяют четыре основные модели базы знаний:

- 1) логическая модель (основанная на формулах);
- 2) продукционная модель (основанная на правилах);
- 3) фреймы (фрейм - это минимально возможное описание сущности объекта);
- 4) семантическая сеть (ориентированный граф, отражающий понятия и их отношения) [2].

Модель, основанная на семантической сети, наиболее соответствует современным представлениям об организации долговременной памяти человека [3].

Типы связей, используемые в семантической сети, выбираются в зависимости от решаемой задачи. В реальном мире число связей между понятиями огромно [4]. К основным типам отношений, которые наиболее часто применяются в семантических сетях, можно отнести следующие:

- 1) класс - элемент класса (например, “автомобиль” – “Волга”);
- 2) свойство – значение (“цвет” – “красный”) [3].

Существует модель семантической сети, которая позволяет находить ответ на вопрос [5].

Местоимением называют самостоятельную часть речи, которая используется вместо существительных, прилагательных, числительных и наречий (или их характеристик), чтобы указать на эти существительные, прилагательные, числительные и наречия (а также их признаки и количество), не называя их.

Личные местоимения указывают на объект речи – лицо, о котором идет речь. Местоимения 1 (я/мы) и 2 (ты/вы) лица обозначают участников речи. Местоимения 3 лица (он, она, оно / они) указывают на лиц, которые в речи не участвуют [6].

Таким образом, местоимения в тексте выполняют вспомогательную функцию, указывая на другие слова, и сами по себе не означают конкретное понятие. Поэтому имеет смысл при построении сети заменять местоимения словами, на которые они указывают. Это позволяет не увеличивать размер сети за счёт местоимений и упростить работу с сетью.

**Алгоритм генерации сети для текста с обработкой личных местоимений 3 лица (он, она, оно, они)** может быть представлен следующим образом:

- 1) Перейти к очередному предложению текста (начиная с первого).
- 2) Найти в предложении существительное в начальной форме. Если найдено, выбрать его в качестве подлежащего.
- 3) Если существительное было найдено, то определить его тип (род и число). Сохранить соответствие данного типа к найденному существительному.
- 4) Найти в предложении личное местоимение 3 лица (он, она, оно, они).
- 5) Если личное местоимение было найдено, то загрузить последнее существительное, соответствующее ему по типу (род и число). Выбрать загруженное существительное в качестве подлежащего.
- 6) Найти сказуемое. Добавить его в сеть, если ещё не добавлено.
- 7) Добавить связь «подлежащее-сказуемое».

8) Добавить остальные члены предложения и связи в соответствии с выбранными для сети типами связей.

9) Повторять шаги 1-8 для каждого предложения в тексте.

Для определения членов предложения в *простейшем* случае могут использоваться словари (существительных, глаголов, прилагательных и т.п.). Для определения рода и числа в *простейшем* случае можно анализировать окончания.

Рассмотрим построение семантической сети на конкретном примере.

Пусть в сети используются типы связей «подлежащее-сказуемое» и «место». Для текста «*Книга повествует о Телемаке, сыне Одиссея. Она лежит на столе.*» можно построить сеть, изображённую на рисунке 1.

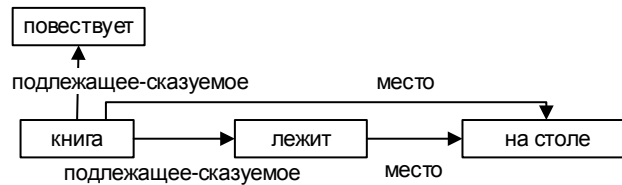


Рисунок 1 – Пример семантической сети

В соответствии с алгоритмом, в сеть последовательно добавлены узлы «книга», «повествует», «лежит», «на столе». Следует отметить, что существительное «книга» имеет женский род и единственное число. Местоимение «она» соответствует данному существительному. Остальные слова не имеют типов связей, для которых строится сеть. Поэтому они в сеть не добавлены.

В исходном тексте не было напрямую сказано, что книга находится на столе, но полученная сеть содержит эту информацию благодаря верному учёту местоимения «она».

Итак, местоимение является частью речи, которая используется довольно часто. Построение семантической сети, учитывающее значение личных местоимений 3 лица, может быть автоматизировано. Тем не менее, реализация алгоритма построения семантической сети представляет собой довольно сложную задачу из-за неоднозначности слов естественного языка.

#### Список использованных источников:

1. Базы знаний экспертных систем. [Электронный ресурс]. – Электронные данные. – Режим доступа: <http://daxnow.narod.ru/index/0-18>. Дата доступа : 24.03.2019.
2. Гаврилова, Т.А. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский – СПб. : Питер, 2000. – 384 с.
3. Масленникова, О.Е. Основы искусственного интеллекта [Электронный ресурс] : учеб. пособие / О.Е. Масленникова, И.В. Гаврилова. – М.: ФЛИНТА, 2013. – Режим доступа: <http://search.rsl.ru/ru/record/01007574162>.
4. Рахимова Д. Р. Построение семантических отношений в машинном переводе // Вестник КазНУ им. аль-Фараби. Серия «Математика, механика и информатика». – Алматы, 2014. – №1. – С. 90-101.
5. Потараев, В.В. Алгоритм применения семантической сети для поиска ответа на вопрос / В.В. Потараев // Компьютерные системы и сети: материалы 54-ой научной конференции аспирантов, магистрантов и студентов. – Минск: БГУИР, 2018. – с. 103-105.
6. Все, что вы хотели узнать о местоимениях, но не знали, как спросить [Электронный ресурс]. – Электронные данные. – Режим доступа: <https://blog.tutoronline.ru/vse-cto-vy-hoteli-uznat-o-mestoimenijah-no-ne-znali-kak-sprosit>. Дата доступа : 24.03.2019.

## ВЕБ-ПРИЛОЖЕНИЕ УЧЕТА РАБОЧЕГО ВРЕМЕНИ ПРЕПОДАВАТЕЛЯ

Рогачевский М. В.

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

Прохорчик Р. В. – ст. преп., м. т. н.

Время является одним из основных ресурсов организации труда на предприятии. Используя время в качестве численной меры труда, можно составлять планы выполнения работ и защитить предприятие и рабочих от большого объема рисков, связанных со сроками. В то время как коммерческое предприятие рискует потерять деньги в случае несоблюдения сроков выполнения работ, в учреждениях образования риски другого рода. При некачественном планировании работ в образовательном учреждении студенты и преподаватели рискуют в определенный момент столкнуться с критическим уровнем нагрузки, который пагубно сказывается на качестве образовательного процесса.