

## ПРОЕКТИРОВАНИЕ БАЗЫ ЗНАНИЙ РЫНКА НЕДВИЖИМОСТИ НА ОСНОВЕ АППАРАТА СЕМАНТИЧЕСКИХ СЕТЕЙ

Руднев Д. К.

*Белгородский государственный национальный исследовательский университет*

*E-mail: 1263800@bsu.edu.ru*

Усовершенствование информационных технологий в современном мире происходит уже порядка 25 лет. И в современном обществе уже тяжело представить, как обходился бы современный человек без средств технического обеспечения, информационно-коммуникационных технологий и персонального компьютера в своей жизни и профессиональной сфере деятельности. Информация стала одним из наиболее значимых ресурсов, а информационные системы являются необходимым инструментом в большинстве профессиональных областях. Современный человек настолько часто сталкивается со всеобщей информацией, что для её обработки требуется разработка большого количества хранилищ разнообразного назначения. Выбранная предметная область имеет обширное число понятий и связей между ними, поэтому объектом исследования была выбрана семантическая сеть. Она позволяет в одинаковой мере управлять как данными, так и знаниями о них. Статья посвящена задаче проектирования базы знаний агентства недвижимости на основе аппарата семантических сетей. Рассмотрены основные особенности и понятия предметной области, что дало возможность определить основополагающие элементы семантической сети, необходимые для решения поставленной задачи. На основе проведенных исследований разработана структура базы знаний, с целью повышения степени наглядности семантического описания процессов агентства недвижимости.

**Ключевые слова:** рынок недвижимости, семантическая сеть, база знаний.

## DESIGNING THE KNOWLEDGE BASE OF THE REAL ESTATE MARKET BASED ON THE SEMANTIC NETWORK APPARATUS

Rudnev D. K.

The improvement of information technologies in the modern world has been going on for about 25 years. And in modern society it is already hard to imagine how a modern person would manage without means of technical support, information and communication technologies and a personal computer in his life and professional field of activity. Information has become one of the most significant resources, and information systems are a necessary tool in most professional areas. Modern man is so often confronted with comprehensive information that its processing requires the development of a large number of repositories for various purposes. The selected domain has a vast number of concepts and links between them, so the semantic network was chosen as the object of study. It allows you to equally manage both data and knowledge about them. The article is devoted to the task of designing a knowledge base of a real estate agency based on the apparatus of semantic networks. The main features and concepts of the subject area are considered, which made it possible to determine the fundamental elements of the semantic network necessary to solve the problem. On the basis of the research conducted, a knowledge base structure has been developed in order to increase the visibility of the semantic description of the real estate agency's processes.

**Keywords:** real estate market, semantic network, knowledge base.

Современный уровень развития бизнеса и сложность задач, стоящих перед руководством, значительно увеличивают важность активного использования современных информационных и коммуникационных технологий в задачах принятия решений. Программные и аппаратные решения сегодня могут предоставить пользователю эффективный

и надежный сбор, хранение, обработку и целевое распределение данных среди различных участников процедур принятия решений.

Осознание этого факта дало толчок к разработке специализированных программных систем, каждая из которых является экспертом в какой-то довольно узкой предметной области, которые называются экспертными системами. Логическая обработка информации с целью получения новых данных, которые явно в нее не вводились, является основной функцией экспертных систем. Такие системы позволяют пользователям, не имеющим навыков программирования, создавать практически значимые приложения, что резко расширяет область применения компьютерных технологий. При решении практических задач экспертные системы позволяют получать результаты, которые сопоставимы, а иногда даже превосходят результаты, которые может получить эксперт [3].

В информационных технологиях концепция экспертных систем неразрывно связана с базами знаний. В результате они рассматриваются как модели поведения экспертов в определенной области знаний с использованием процедур логического вывода и принятия решений, а база знаний – как совокупность фактов и правил логического вывода в конкретной предметной области [5].

При разработке и проектировании семантической сети не используются ограничения на количество элементов и связей. Поэтому систематизация отношений и связей между объектами в сети необходима для дальнейшей формализации.

Эффективность экспертной системы в основном определяется объемом и качеством содержащихся в них формализованных экспертных знаний, которые представлены как в форме теоретических, так и практических навыков [4].

База знаний должна содержать данные, необходимые агентам, работающим с семантической памятью, для организации коллективных действий для решения проблем, с которыми должна справляться экспертная система. Структурирование является необходимым шагом в разработке модели семантической базы знаний для расширения разнообразия типов определений и концепций, которые хранятся в базе данных.

Понятие базы знаний тесно связано с понятием предметной области. Отношение между базой знаний и предметной областью, которую она описывает, определяет семантику базы знаний интеллектуальной системы. Семантические сети представляют собой формализованное представление знаний в форме ориентированного графа, где множество вершин формируется объектами или ситуациями, а множество дуг формируется отношениями

между ними. Выводы по семантическим сетям определяются отношением между множеством дуг, имеющих общие узлы. Если вершины не имеют своей собственной структуры, то соответствующие сети называются простыми сетями. И наоборот, если вершины имеют определенную структуру, то такие сети называются иерархическими [1].

Одним из основных отличий иерархических семантических сетей от простых семантических сетей является возможность делить сеть на подсети (пространства) и устанавливать отношения не только между вершинами, но и между пространствами. Все вершины и дуги являются элементами хотя бы одного пространства. Понятие пространства аналогично понятию скобок в математической нотации. Различные пространства, являющиеся частью сети, могут быть организованы в виде дерева пространств, вершины которых соответствуют пространствам, а дуги являются отношениями видимости. На рис. 1 приведен пример дерева пространств, в соответствии с которым, например, из пространства  $P_6$  (пространство-потомок) есть возможность увидеть все вершины и дуги, которые лежат в пространствах-предках  $P_4$ ,  $P_2$ ,  $P_0$ , а остальные пространства невидимы. Отношение видимости дает возможность структуризации пространства в упорядоченные множества, называемые перспективами. Перспектива часто используется для ограничения сетевых объектов, которые видимы для некоторых процедур, работающих с сетью. Обычно перспектива включает в себя не какие-либо, а иерархически структурированные пространства.

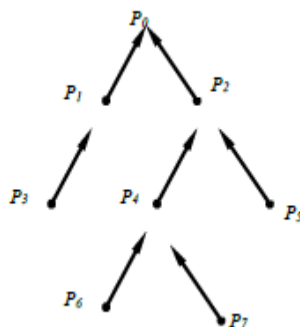


Рисунок 1. Пример разбиения сети на пространства

Математические методы позволяют описывать большое количество явлений в окружающем нас мире в форме логических высказываний. Семантические сети возникли как попытка визуализировать математические формулы. Основным представлением для семантической сети является граф. Но не стоит забывать, что за графическим отображением непременно стоит строгая математическая запись и что обе эти формы отображают одно и то же. В математике граф представляется набором вершин  $V$  и множеством отношений между

ними Е. Используя аппарат математической логики, можно сделать вывод, что каждая вершина соответствует элементу предметного множества, а дуга – предикату.

Таким образом, информацию о предложениях рынка недвижимости может быть представлена с помощью математической формализации в виде кортежа:

$$K = (I, L, F, C_I, T_I, T_L, M_I, M_L, M_{IL}) \quad (1)$$

В формуле (1):

$I$  – множество узлов семантической сети;

$L$  – множество дуг семантической сети;

$F(I, L)$  – функция инцидентности узлов и дуг семантической сети;

$C_I(I)$  – функция, возвращающая информационное наполнение узлов семантической сети;

$T_I(I)$  – функция, которая возвращает тип узлов семантической сети;

$T_L(L)$  – функция, которая возвращает тип (окраску) дуг семантической сети;

$M_I$  – множество возможных типов узлов семантической сети;

$M_L$  – множество возможных типов (окрасок) дуг семантической сети;

$M_{IL}$  – множество троек вида  $(I, J, L)$ , где  $I \in T_I, J \in T_I, L \in T_L$ . Если  $(I, J, L) \in M_{IL}$ , то это значит, что узлы с типами  $I$  и  $J$  могут быть связаны ориентированной связью типа  $L$ .

Компоненты  $I, L, F(I, L), C_I(I), T_I(I)$  и  $T_L(L)$  содержат информационное наполнение семантической сети, которое при фиксированных  $I, L, F(I, L), T_I(I)$  и  $T_L(L)$  представимо в виде вектора конечной длины. Множества  $M_I, M_L$  и  $M_{IL}$  содержат описание предметной области данных и задаются во время начальной настройки экспертной системы.

Семантические сети систематизируют функции отношений между понятиями посредством следующих функций:

- множество – подмножество (типы отношений «абстрактный - конкретный», «целое - часть», «род - тип»);

- индексы (свойства, прилагательные в языке и т.п.);
- отношения «близость»;
- отношения «сходства – различия»;
- отношения «причина – следствие»;
- связи, которые определяют тип объектов ("это есть" либо "класс-подкласс", "иметь частью" либо "часть-целое", "принадлежать" либо "элемент множество" и т.п.);
- функциональные отношения (обычно определяются глаголами «производит», «влияет»);
- количественные ("больше", "меньше", "равно");
- пространственные ("далеко от", "близко от", "за", "под", "над");
- временные ("раньше", "позже", "в течение");
- атрибутивные отношения (иметь свойство, иметь значение);
- логические связи ("и", "или", "не") и др.

Понятия семантической сети записываются в овалах или прямоугольниках и связаны стрелками с подписями – дугами. Эта форма наиболее удобна для восприятия человеком. Однако в семантических сетях есть недостатки при разработке более сложных сетей или при попытке учесть особенности естественного языка. Структуры семантических сетей, где указываются направления навигационных отношений, называются картами знаний, а их комбинация, позволяющая охватить большие области семантической сети, представляет собой атлас знаний [6].

Для разработки сетевой модели представления знаний, необходимо выполнить следующие шаги [2]:

- 1) Определение абстрактных объектов и понятий предметной области, необходимых для решения конкретной поставленной задачи. Оформить их в виде вершин.

- 2) Задать свойства для выделенных вершин, оформив их в виде вершин, связанных с исходными вершинами атрибутивными отношениями.
- 3) Задать связи между этими вершинами, с помощью функциональных, пространственных, количественных, логических, временных, атрибутивных отношений, а также отношений типа «являться наследником» и «являться частью».
- 4) Добавление конкретных объектов и понятий, описывающих решаемую задачу. Оформить их в виде вершин, которые связаны с уже существующими отношениями типа «являться экземпляром», «есть», is\_a.
- 5) Проверка правильности установленных отношений по следующему принципу: вершины и само отношение при правильном построении образуют предложение.

Декларативные и процедурные знания в семантических сетях не разделены, поэтому база знаний не отделена от механизма логического вывода. Процедура вывода часто представляет собой набор процедур сетевой обработки.

Задача поиска решения в базе знаний типа семантической сети сводится к задаче поиска фрагмента сети, который соответствует определенной подсети, отвечающей на поставленный вопрос. Подобные проблемы решаются с помощью инструмента теории графов. Фундаментальные атрибуты соединений, такие как: рефлексивность, симметрия и транзитивность, играют важную роль в процессе вывода в сеть.

Для достижения поставленной цели были предприняты следующие шаги:

- 1) Ключевые понятия данной предметной области – агентство недвижимости, тот, кто посещает агентство недвижимости (клиент) и те, кто его обслуживают (риелторы). У риелторов и клиентов есть общие характеристики, поэтому целесообразно выделить общее абстрактное понятие – человек. Предоставляемыми услугами агентства недвижимости являются сделки с объектами недвижимости, которые хотят совершить клиенты. Исходя из этого, вершины графа будут следующими: «Агентство недвижимости», «Человек», «Риелтор», «Клиент», «Сделка» и «Объект».
- 2) У этих объектов есть определенные свойства и атрибуты. Например, агентства недвижимости и объекты располагаются по определенным адресам, каждый объект имеет

свои особенности: количество комнат, площадь, цену и т.д. Поэтому добавим вершины «Адрес», «Цена», «Комнаты» и «Площадь».

3) Зададим связи между этими вершинами, используя функциональные, пространственные, количественные, логические, временные, атрибутивные отношения, а также отношения типа «являться наследником» и «являться частью».

4) Добавление знаний о конкретных фактах решаемой задачи. Пусть имеется агентство недвижимости «Грант», в котором работает риелтор Сергей. Потенциальный клиент Вадим решил пойти в агентство недвижимости и выбрать однокомнатную квартиру для съема за 10000 рублей. Также известны адреса агентства и объекта.

5) Осуществление проверки установленных связей. Например, возьмем вершину «Объект» и пройдем по установленным связям. На выходе получаем следующую информацию: объект является частью сделки, примерами объектов могут служить квартира, которая стоит 10000 рублей.

Структура сематической сети для заданной предметной области представлена на рисунке 2.

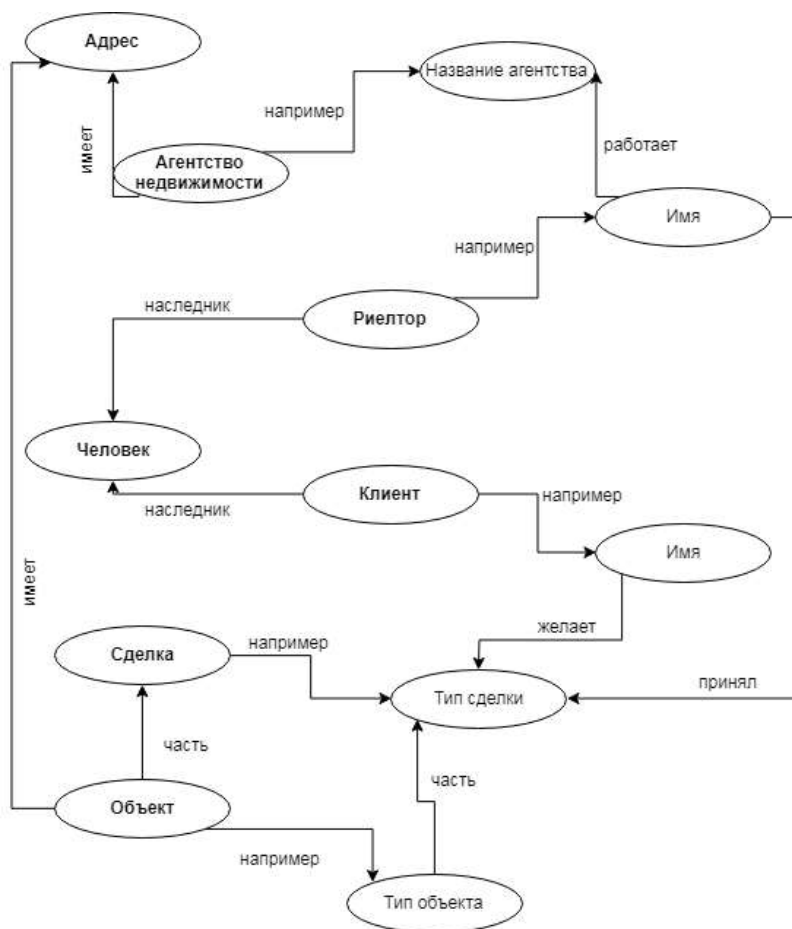


Рисунок 2. Структура базы знаний на основе семантической сети

В ходе написания статьи была детально изучена и исследована предметная область агентства недвижимости, его структуры и процессы. В работе показаны основные особенности и понятия предметной области, что позволило определить необходимые элементы семантической сети, на основе которых была разработана структура базы знаний агентства недвижимости.

### Список литературы

1. Асадуллаев Р.Г. Базы знаний и экспертные системы: учебное пособие / Р.Г. Асадуллаев, Т.А. Лысакова; рец.: А.А. Черноморец, А.Г. Бажанов; М-во образования и науки РФ, НИУ БелГУ. – Белгород: ИД Белгород, 2018. – 176 с.
2. Джарратано Д. Экспертные системы: принципы разработки и программирование. 4-е изд.: учебник / Д. Джарратано, Г. Райли. – М.: "Вильямс", 2007. – 1152 с.
3. Зайцева Т.В. Интеллектуальные информационные системы: Электронный ресурс: Учебно-методический комплекс / Т.В. Зайцева. –Белгород, 2013. – 152 с.
4. Козин Р.Г. Экспертные системы: Учебное пособие/ Р.Г. Козин. – М.: МИФИ, 2008. – 87 с.
5. Ломакин В.В. Базы данных и базы знаний: Электронный ресурс: Учебное пособие / В.В. Ломакин. – Белгород: БелГУ, 2010. – 216 с.
6. Хижняков Д.П. Организация работы с базами знаний и экспертными системами: Электронный ресурс: Учебно-методический комплекс / Д.П. Хижняков. – Белгород, 2013. – 148 с.