



УДК 004.822:514

ВИРТУАЛЬНЫЙ ПОМОЩНИК АБИТУРИЕНТА ПРИ ВЫБОРЕ СПЕЦИАЛЬНОСТИ

Барлыбаев А.Б., Нургазинова Г.Ш.

*Евразийский национальный университет имени Л.Н.Гумилева НИИ «Искусственный интеллект»,
г. Астана, Республика Казахстан*

frank-ab@mail.ru

Цель работы – автоматизация формирования ответа по запросу пользователя к базе знаний заданной предметной области. Предметная область – абитуриент выбирает специальность. База знаний предметной области представляется в виде семантической сети, состоящей из фактов и правил вывода. Требуется построить две семантические сети. Первая сеть строится экспертом, а вторая – абитуриентом, отвечая на вопросы, которые представлены на своих пользовательских интерфейсах. При этом функция решателя задач интеллектуальной системы сводится к выяснению вопроса «является ли семантическая сеть абитуриента частью семантической сети эксперта». Если да, то формируется ответ абитуриенту о подходящей ему специальности.

Ключевые слова: семантические сети, решатель задач, виртуальный помощник абитуриента, e-learning.

Введение

Семантическая сеть – это взвешенный ориентированный граф, в котором весами вершин являются понятия, события, объекты или субъекты, а весами дуг – отношения между вершинами. Любой фрагмент сети, например, одна вершина, две вершины и дуги между ними называют подсетью. Логический вывод (поиск решения) на семантической сети заключается в том, чтобы найти или сконструировать подсеть, удовлетворяющую некоторым условиям [Roussopoulos, 1976] [Quillian, 1968] [Sowa, 1987].

Отношения, представляемые дугами, в семантической сети могут быть различными. Примеры таких отношений приведены в таблице 1.

Типы отношений выбираются в зависимости от вида семантической сети (таблица 2) и решаемой задачи.

Таблица 1. Основные виды отношений в семантических сетях.

Тип	Описание
Является наследником (a-kind-of)	Задаёт иерархические связи между классами
Является экземпляром (is-a)	Определяет значение, описывает конкретный объект, понятие
Это есть	Может использоваться

(are)	вместо связи a-kind-of в отношениях подразумевающих равенство или эквивалентность
Быть частью (has-part)	Определяет структурные связи, описывает части или целые объекты
Функциональные или предикативные	Определяются обычно глаголами, отражают различные действия или отношения (учить, владеть и т.д.).
Количественные	Отображают количественные соотношения между вершинами (больше, меньше и т.д.)
Пространственные	Отображают пространственные отношения между вершинами (близко, далеко и т.д.)
Временные	Описывают временные связи между вершинами (скоро, долго, сейчас и т.д.)
Атрибутивные	Описывают свойства объектов, понятий
Логические	Описывают логические связи между вершинами (и, или, не)

Здесь глаголы могут быть разного падежа в зависимости от типов вершин и вида направленности отношения между ними, например: Субъект – учитель, субъект – ученик. «Учитель *учит* ученика», либо «ученик *учится* у учителя».

Таблица 2. Типы семантических сетей.

Тип	Описание
По типу знания	
Экстенциональные (терминальные)	Описывает конкретные отношения над конкретными значениями.
Интенциональные (нетерминальные)	Описывают имена понятий, событий, объектов и субъектов, а не их индивидуальные имена или значения, и отношения между ними.
По типу ограничений на вершины и дуги	
Простые	Вершины сети не обладают внутренней структурой (терминальные)
Иерархические	Вершины сети обладают внутренней структурой (нетерминальные).
Динамические (сценарии)	Вершины сети представляют некоторые действия или события
По количеству типов отношений	
Однородные	Обладают только одним типом отношений
Неоднородные	Количество типов отношений больше двух
По арности отношений	
Двуместные (Бинарные)	Все отношения в сети связывают ровно две вершины
Многместные (n -арные, где $n > 2$ - число вершин)	В сети есть отношения, связывающие более двух объектов

Здесь *события* представляют действия, происходящие в реальном мире, и определяются указанием типа действия и ролей, которые играют объекты в этом действии.

Заметим, что в иерархической сети есть возможность разделять сеть на подсети и устанавливать отношения не только между вершинами, но и между подсетями (различные подсети, существующие в сети, могут быть упорядочены в виде дерева подсетей, вершины которого - подсети, а дуги - отношения видимости).

Важную роль в построении семантической сети играет предметная область, которая определяется следующим образом:

1. Определить абстрактные объекты и понятия предметной области, необходимые для решения 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, используя функциональные, пространственные, количественные, логические, временные, атрибутивные отношения, а также отношения типа «являться наследником» и «являться частью».

4. Добавить конкретные объекты и понятия, описывающие решаемую задачу. Оформить их в виде вершин, связанных с уже существующими отношениями типа «являться экземпляром», «есть».

5. Проверить правильность установленных отношений (вершины и само отношение при правильном построении образуют предложение, например «Двигатель является частью автомобиля»).

1. Проектирование виртуального помощника с помощью семантической сети

Задача 1. Построить сетевую модель представления знаний в предметной области «Абитуриент выбирает специальность». Выполняя вышеуказанные пункты 1-5, эту предметную область можно определить как:

1. Ключевые понятия данной предметной области – специальность, абитуриент. Абитуриент выбирает специальность для поступления. Специальность является частью группы специальностей. Группа специальностей часть уровней образования. При выборе специальности абитуриент использует характеристики имеющее на данный момент: образование и позицию (работу) на которой он хотел бы трудиться. Нам необходимо определить к какой части сектора труда относится позиция. В зависимости какое образование есть у абитуриента на данный момент, важную роль играет 4-тый предмет, сдаваемый при вступительных экзаменах (единое национальное тестирование). 4-тый предмет часть среднего образования. Важную роль играет также имеющийся у абитуриента уровень высшего образования и специальность. Исходя из этого, вершины графа будут следующими: «Абитуриент», «Специальность», «Образование», «Сектор труда», «Позиция», «4-тый предмет», «Уровень образования», «Уровень образования и специальность» и «Группа специальностей».

2. У этих объектов есть определенные свойства и атрибуты. Например, образование может

быть средним или высшим, уровень образования может быть бакалавриатом, магистратурой или докторантурой PhD. Поэтому добавим вершины «Среднее образование», «Высшее образование», «Бакалавриат», «Магистратура» и «Докторантура PhD».

3. Определим для имеющихся вершин отношения и их типы, используя таблицу 2. Полученный в результате граф изображен на рис. 1.

4. Добавим знание о конкретных фактах решаемой задачи. Пусть имеется два абитуриента: «Марк» и «Алекс», у первого есть только среднее образование, а у второго высшее образование. Марк при сдаче Единого Национального Тестирования выбрал 4-тым предметом физику. Алекс имеет высшее образование бакалавр экономики. Есть две группы специальностей «Социальные науки, экономика и бизнес» и «Естественные науки». У первой группы специальностей есть специальности «менеджмент» и «учет и аудит», у второй только «информатика». Каждая группа специальностей имеет весь набор уровней образования: бакалавриат, магистратура, докторантура PhD. Марк хочет работать в будущем программистом. А Алекс хочет быть топ-менеджером. Исходя из этого, добавим соответствующие вершины в граф и соединим их функциональными отношениями и отношениями типа «например или является экземпляром». Стоит учесть, что отношения «например» и «является экземпляром» эквивалентны. Полученный в результате граф изображен на рис. 2.

5. Осуществим проверку установленных связей. Например, возьмем вершину «Специальность» и пройдем по установленным связям. Получаем следующую информацию: специальность является частью группы специальности, примерами специальности могут служить менеджмент, учет и аудит, информатика.

Задача 2. Что должен делать абитуриент, чтобы выбрать нужную специальность. Для решения этой задачи нужно использовать Правила приема в высшие учебные заведения Республики Казахстан [Правительство РК, 2012].

Следуя правилам, абитуриент с средним образованием может поступать в высшие учебные заведения на специальности бакалавриата в зависимости от 4-го предмета, определяющего направление обучения (группу специальностей). А абитуриент, с высшим образованием, может поступать на все направления специальностей в бакалавриате. Для поступления на специальность магистратуры необходимо иметь образование бакалавриата по данному направлению обучения (группа специальности). Для поступления на специальность докторантуры PhD необходимо иметь образование магистратуры по данному направлению обучения (группа специальности).

Также абитуриенту необходимо определить

сектор труда и будущую профессию, по которой он в дальнейшем будет трудиться.

Чтобы выбрать правильную специальность, учитывая выше упомянутые факторы, абитуриент заполняет анкету-форму. Абитуриент отвечает на вопросы:

1. Вопрос «какое имеет образование?»
Варианты ответа: «среднее», «высшее».

1.1 Если выбрано «среднее», то вопрос «какой 4-тый предмет?»

1.2 Если выбрано «высшее», то вопрос «какой уровень образования и какая специальность?»

2. Вопрос «какой сектор труда?» и «какая позиция?»

На основе ответов строится семантическая сеть «анкета-форма». Далее полученные данные передаются в решатель задач.

Решатель задачи – процедура, которая выясняет, является ли построенная сеть «анкета-форма» частью «Абитуриент выбирает специальность».

Для получения ответа на какой-либо вопрос по этой задаче, необходимо найти соответствующий участок сети и, используя связи, получить результат.

Например, вопрос «Имея высшее образование бакалавр экономики, и имея желание работать топ менеджером, какую специальность должен выбрать Алекс?» Из запроса понятно, что необходимо найти следующие вершины: «Алекс», «Специальность», «Высшее образование», «Бакалавр экономики» и «Топ-менеджер». Часть семантической сети, находящаяся между этими вершинами, содержит ответ, а именно, высшее образование бакалавр экономики дает возможность поступать на все специальности бакалавриата, и на специальности магистратуры социальных наук и экономики и бизнеса. Соответственно, Алекс может поступить на бакалавриат специальностей информатики, учет и аудит, менеджмент, а также на магистратуру учет и аудит, менеджмент. Но желаемая позиция связана с одной специальностью, менеджментом. Больше информации в модели нет, поэтому делаем вывод – Алекс может поступить на бакалавриат или магистратуру специальности менеджмент.

Задача 3. Описание работы «виртуального помощника абитуриента при выборе специальности». На рис. 3 представлена блок-схема работы виртуального помощника.



Рис. 1. Интенциональная (нетерминальная) семантическая сеть предметной области «Абитуриент выбирает специальность».



Рис. 2. Экстенциональная (терминальная) семантическая сеть предметной области «Абитуриент выбирает специальность».

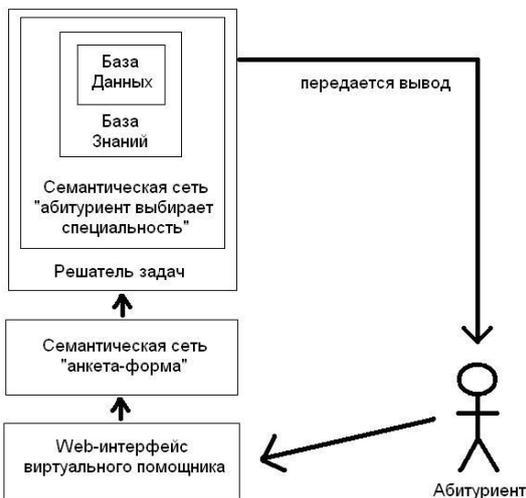


Рис. 3. Блок-схема виртуального помощника.

Абитуриент с помощью веб-интерфейса виртуального помощника строит свою семантическую сеть «анкета-форма». Построенная модель передается в решатель задач. Решатель задач состоит:

1. База данных, атомарный блок, в котором хранятся данные, факты.
2. База знаний, блок в котором хранятся правила обработки данных, фактов.
3. Семантическая сеть «абитуриент выбирает специальность», строится на основе базы данных и базы знаний.

Решатель задач сравнивает семантические сети «анкета-форма» и «абитуриент выбирает специальность». Если решатель задач выясняет, что построенная сеть «анкета-форма» является частью «абитуриент выбирает специальность», то запускается поиск решения по семантической сети. Решением будет промежуточные вершины, которые зависят от семантической сети «анкета-форма», но при этом существуют в семантической сети «абитуриент выбирает специальность». Подробнее описано в рис. 4.

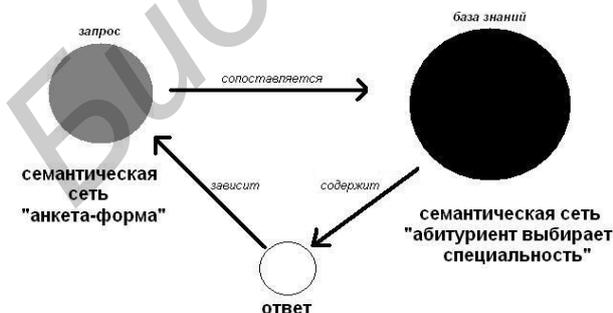


Рис. 4. Визуализация поиска решения в семантической сети.

Ниже будет построен алгоритм решателя задач, который сравнивает семантические сети «анкета-форма» и «абитуриент выбирает специальность».

Опишем следующие объекты, которые будут

использоваться при построении данного алгоритма в таблице 3.

Таблица 3. Описание объектов семантических сетей, используемых в решателе задач.

№	Обозначение объекта	Описание
1	->	Оператор связанно
2	∈	Оператор содержит
3	E	Образование
4	HE	Высшее образование
5	GS	Группа специальности
6	LEaS	Уровень образования и специальность
7	LE	Уровень образования
8	AGS	Список групп специальностей абитуриента
9	LM	Сектор труда
10	P	Позиция
11	ALM	Сектор труда абитуриента
12	AP	Позиция абитуриента
13	S	Специальность
14	AS	Список подходящих специальностей для абитуриента
15	FS	Четвертый предмет
16	AFS	Четвертый предмет абитуриента

НАЧАЛО АЛГОРИТМА «Решатель задач»

Исходные данные:

1. Семантическая сеть абитуриента анкета-форма, которая состоит из объектов (вершин): LEaS, ALM, AP, AFS;
2. Семантическая сеть эксперта абитуриент выбирает специальность, которая состоит из объектов (вершин): E, HE, GS, LE, LM, P, S, FS;

Результирующие данные:

1. Да: AS;
2. Нет: выполняется команда, выводящая сообщение «просим прощения, но у нас нет подходящей специальности для вас»; выполняется команда quit 0, закрывающая действие алгоритма;

Метод:

Сравнение семантической сети абитуриента с семантической сетью эксперта

Проверка на имеющиеся образование

Если высшее образование, То формируется AGS в зависимости от LEaS

Иначе формируется AGS в зависимости от AFS

Фильтрация AGS в зависимости от ALM и AP

AGS ∈ GS -> S

Если S -> P = AP and P -> LM = ALM, То AS = S

Если они совпадают или AS' = "", То ДА

Иначе НЕТ

КОНЕЦ АЛГОРИТМА «Решатель задач»

Этот алгоритм был реализован на языке

программирования *Cache Object Script (Mumps)* с использованием технологии *постреляционной СУБД Cache Intersystems* [Intersystems, 2014].

Заключение

Даны определения семантической сети. Спроектирована работа виртуального помощника абитуриента при выборе специальности. При проектировании были использованы методы нахождения информации в семантических сетях. Была спроектирована работа и структура решателя задач в семантических сетях. Построен алгоритм решателя задач, который сравнивает семантическую сеть эксперта с семантической сетью абитуриента. Написана программная реализация данного алгоритма. Проведены начальные экспериментальные наблюдения работы алгоритма. Данная проблема может быть расширена до решения задачи улучшения эффективности системы образования с учетом региональных трудовых ресурсов.

Библиографический список

[Roussopoulos, 1976] Roussopoulos N.D.A semantic network model of data bases. – TR No 104, Department of Computer Science, University of Toronto, 1976.

[Quillian, 1968] Quillian, M. R. (1968). Semantic memory. Semantic information processing, 227–270.

[Sowa, 1987] John F. Sowa (1987). "Semantic Networks". In Stuart C Shapiro. Encyclopedia of Artificial Intelligence. Retrieved 2008-04-29.

[Правительство РК, 2012] Типовые правила приема на обучение в организации образования, реализующие профессиональные учебные программы высшего образования. Утверждены постановлением Правительства Республики Казахстан от «19» января 2012 года № 111 (с изменениями от 19 апреля 2012 года № 487)

[Intersystems, 2014] <http://www.intersystems.com/cache/>

THE VIRTUAL ASSISTANT FOR APPLICANT IN CHOOSING OF THE SPECIALTY

Barlyabyev A.B. *, Nurgazinova G.Sh. *

* *L.N.Gumilyov Eurasian National University, SRI
«Artificial Intelligence»,
Astana, Republic of Kazakhstan*

frank-ab@ mail.ru

This article's purpose is automation of answer formation to the user's request to the domain's knowledge base. The domain –the applicant chooses a specialty. Domain knowledge base is represented as a semantic network consisting of facts and rules of conclusion.

Introduction

The semantic network is the weighted oriented graph, in which vertex weights are concepts, events, objects or entities, and weights of arcs - the relation between the vertices. Any fragment of network, such as one vertex, two vertices and arcs between them is called subnet. The logical conclusion (decision search) on the

semantic network consists in finding or designing a subnet meeting some conditions.

Main Part

START OF THE ALGORITHM “problem solver”

Background:

The applicant's semantic network the profile-form which consists of objects (vertices): LEaS, ALM, AP, AFS;

The expert's semantic network the applicant chooses a specialty which consists of objects (vertices): E, HE, GS, LE, LM, P, S, FS;

The resulting data:

Yes: AS;

No: there are executing the command that displays the message “Sorry, but we do not have a suitable specialty for you”, quit 0 command is executed, this is stop operation of the algorithm;

Method:

Comparison of applicant's semantic network with a semantic network of the expert

Check on available education

If higher education, Then AGS is formed depending on LEaS

Else, AGS is formed depending on AFS

AGS filtration depending on ALM and AP

AGS ∈ AGS → S

If S → P = AP and P → LM = ALM, then AS = S

If they are identical or AS' = “”, Then YES

Else NO

END OF THE ALGORITHM “problem solver”

This algorithm has been implemented in the Cache Object Script (Mumps) programming language with using of technology of post-relational DBMS Cache Intersystems.

Conclusion

Definitions of the semantic network are given. The applicant's virtual assistant operation in choosing specialty is designed. At designing methods of information searching in semantic networks were used. It was projected operation and structure of the problem solver in semantic networks. The algorithm of a problem solver is constructed which compares the expert's semantic network with the semantic network of the applicant. Program realization of this algorithm is written. The initial experimental observations of the algorithm is conducted. This problem can be expanded up to solve the problems of the improvement the efficiency of the education system according to regional labor resources.