

## МОДЕЛИРОВАНИЕ ЯЗЫКОВОЙ ДЕЯТЕЛЬНОСТИ

УДК 519 766.

### **БАЙМАТОВА Мария-Малика Равшанбековна**

магистр, Днепропетровский национальный университет.

**Научные интересы:** моделирование языковой деятельности методами искусственного интеллекта.

### **МАРАСАНОВ Владимир Васильевич**

д.т.н., профессор, зав. кафедрой технической кибернетики, Херсонский национальный технический университет.

**Научные интересы:** системы искусственного интеллекта, экспертные системы.

### **КОСЕНКО Игорь Александрович**

аспирант кафедры технической кибернетики, Херсонский национальный технический университет.

**Научные интересы:** конечные автоматы в задачах идентификации дискретных систем.

Над решением задач моделирования языковой деятельности в настоящее время работают коллективы исследователей в Европе, США, России. Проблема понимания естественного языка в значительной степени зависит от знания предметной области и является одним из важнейших направлений исследований по искусственному интеллекту. При создании лингвистического анализатора работы исследователей проводятся по направлениям фонологии, морфологии, синтаксического анализа, семантики, контекстно-свободных грамматик.

Одной из важнейших лингвистических задач в области проблематики искусственного интеллекта является построение таких лингвистических процессоров, которые были бы в состоянии обеспечить человеко-машинное общение на естественном языке. Решение этой задачи сводится к разработке адекватной лингвистической модели языкового взаимодействия. Модель языкового взаимодействия может быть сведена к системе автоматов, отражающая коммуникативную среду, языковую и коммуникативную компетенцию партнёров общения. Язык при этом мыслится как важнейший инструмент осуществления рационального поведения человека, в связи с чем устройство языка предопределено устройством сознания и в то же время моделирование сознания в известной степени может осуществляться с использованием языковых свидетельств.

В соответствии с этим принципом взаимодействуют компоненты семантического представления, создаётся и используется арсенал синтаксических средств выражения полипредикативности, строится предикатно-аргументная структура. Этот принцип обладает объяснительной силой при описании многих фрагментов языковой структуры. Более того, лексическое значение слов также сохраняет следы действия принципа приоритета.

При автоматическом анализе и синтезе целого текста, как и в процессе человеко-машинного общения в режиме диалога, наряду с проблемами, связанными с синтаксисом естественного языка или же с его лексикой, вытекают проблемы, связанные с референцией – т. е. соотношением тех или иных языковых сущностей с сущностями внеязыковыми. Для придания лингвистике черт естественной науки в системах искусственного интеллекта используются практически все разделы математики и математической логики.

Графы оказались идеальным средством для формализации ассоционистских теорий знаний за счёт точного представления отношений посредством дуг и узлов. Семантическая сеть представляет знания в виде графа, узлы которого соответствуют факторам или понятиям, а дуги – отношениям или ассоциациям между понятиями. Сетевые представления имеют почти такую же длинную историю, как и логика. Наиболее

ранней работой, имеющей прямое влияние на современные семантические сети, была система экзистенциальных графов, разработанная Чарльзом Пирсом в двадцатом столетии [1]. Теория Пирса использовала всю выразительную силу исчисления предикатов первого порядка с аксиоматической основой и формальными правилами вывода. Графы давно используются в психологии для представления иерархий и наследования свойств [2, 3, 4]. Первые компьютерные реализации семантических сетей были созданы в начале 1960-х для использования автоматического (машинного) перевода [4]. Вслед за первыми работами по созданию схем представления знаний был предложен ряд сетевых языков для моделирования семантики естественного языка. Примером сетевого языка представления являются концептуальные графы [5].

Концептуальный граф – это конечный, связный, двудольный граф. Узлы графа представляют понятия или концептуальные отношения (унарные, бинарные, тернарные, ...) [6]. Чтобы различать типы узлов, узлы понятий представляются прямоугольниками, а концептуальные отношения – эллипсами.

В концептуальных графах узлы понятий представлены либо конкретные, либо абстрактные объекты в мире рассуждений. Конкретные понятия характеризуются нашей способностью сформировать их образ (телефон, трактор). Абстрактные понятия включают такие абстракции как любовь, красота, верность, для которых не существуют образов в нашем воображении.

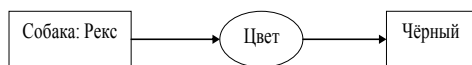


Рисунок 1 – Концептуальный граф, показывающий, что собака Рекс – чёрного цвета

В концептуальных графах каждое понятие является уникальным экземпляром конкретного типа. Каждый прямоугольник понятия снабжается меткой типа, определяющей класс или тип экземпляра, представленного этим узлом [6].

Концептуальные графы позволяют также описывать конкретные, но неименованные экземпляры. Для обозначения каждого экземпляра в мире рассуждений используется уникальный дескриптор, называемый маркером. Он представляет собой число, перед кото-

рым расположен символ #. Маркеры отличаются от имён тем, что они являются уникальными: экземпляр может иметь одно имя, несколько имён или вовсе быть безымянным, но он имеет ровно один маркер. Подобные экземпляры могут быть тезками, но не могут иметь один и тот же маркер. Это различие даёт возможность преодолевать семантические неоднозначности, возникающие при именовании объектов. Граф на рис. 2 означает, что конкретная собака #1352 – чёрная.



Рисунок 2 – Концептуальный граф, показывающий, что отдельная (но безымянная) собака – чёрного цвета

Маркеры позволяют отделить экземпляры от их имён. Если собака #1352 имеет имя Рекс, для его описания можно использовать концептуальное отношение (имя). Результатом является граф Рис. 3. Имя, заключённое в кавычки, указывает, что это строка. Там, где не возникает опасность неоднозначности, мы можем упростить граф и ссылаться на экземпляр прямо по имени. При таком соглашении граф на рис. 3 эквивалентен графу на рис. 1.

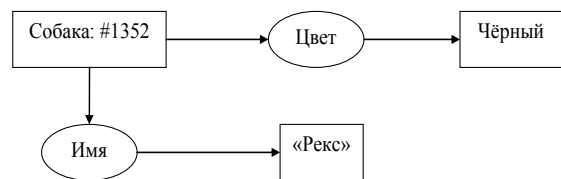


Рисунок 3 – Концептуальный граф, показывающий, что собака с именем Рекс – чёрного цвета

Вместо обращения по маркеру или по имени можно использовать обобщённый маркер \* для определения неспецифицированных экземпляров. Эта информация часто опускается в метках понятий: узел, заданный только меткой типа собака, эквивалентна узлу собака: \*. Дополнительно к обобщённому маркеру концептуальные графы позволяют использовать именованные переменные. Они представляются звёздочкой, за которой следует имя переменной (например, \*x).

Вопросы наследования решаются (частично) с помощью частичного упорядочения на множестве типов, которое можно обозначить символом  $\leq$ . Если  $s$  и  $t$  – есть типы и  $t \leq s$ , то считают, что  $t$  – это подтип, а  $s$  – супертип  $t$ . Поскольку эта систематизация частична, тип может иметь один или несколько супертипов, а также один или несколько подтипов. Если  $s$ ,  $t$  и  $u$  – типы, причём  $t \leq s$  и  $t \leq u$ , то считают, что  $t$  – это общий подтип для  $s$  и  $u$ . Аналогично, если  $s \leq v$  и  $u \leq v$ , то  $v$  – общий супертип для  $s$  и  $u$ . Иерархия типов концептуальных графов представляет собой решётку, описывающую общий вид системы множественного наследования [6].

В решётке типы могут иметь множество родителей и детей. Однако каждая пара типов должна иметь минимальный общий супертип и максимальный общий подтип. Тип  $v$  является минимальным общим супертипом для типов  $s$  и  $u$ , если  $s \leq v$  и  $u \leq v$  и для любого общего супертипа  $w$  типов  $s$  и  $u$   $v \leq w$ . Максимальный общий подтип имеет аналогичное определение. Минимальный общий супертип совокупности типов служит для определения свойств, присущих только этим типам. Поскольку многие типы, такие как «эмоция» и «рок» не имеют общих супертипов или подтипов, необходимо добавить типы, которые выполняют эти роли. Чтобы иерархия типов действительно стала решёткой, в концептуальные графы включают два специальных типа. Универсальный тип обозначается символом  $\top$  и является супертипом всех типов. Абсурдный тип обозначается

символом  $\perp$  и является подтипом всех типов. В теории упорядоченных множеств рис.4 трактуется как диаграмма Хассе [6] и используется при решении задач характеризационного анализа с помощью семантического эквивалентирования.

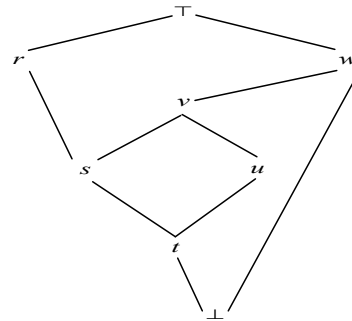


Рисунок 4 – Иерархия типов, иллюстрирующая под-типы, супертипы, универсальный и абсурдный тип. Дуги представляют отношения.

## ВЫВОДЫ

Рассмотрена одна из задач построения систем моделирования языковой деятельности на основе построения и преобразования семантической сети. Проанализирована возможность использования концептуальных графов для построения и преобразования семантических сетей.

## ЛИТЕРАТУРА:

1. Roberts D.D. The Existential Graphs of Charles S. Peirce. The Hague: Mouton, 1973.
2. Collins A. and Quillian M.R. Retrieval time from semantic memory. Journal of Verbal Learning & Verbal Behaviour, 8: 240-247, 1989.
3. Selz. O. Über die Gesetze des Geordneten Denkverlaufs. Stuttgart: Spemann, 1913.
4. Masterman M. Semantic message detection for machine translation, using Interlingua. Proceedings of the 1961 International Conference on Machine Translation, 1961
5. Sowa J. F. Conceptual Structures: Information Processing in Mind and Machine. Reading, M.A. Addison-Wesley, 1984.
6. Gorbatov V.A. Teoriya chastichno uporyadchennykh sistem. M.: «Sovetskoe radio», 1976. – 336 s.
7. Ljagger, Dzhordzh F. Iskusstvennyj intellekt: strategii i metody resheniya slozhnykh problem. – M.: Izdatel'skiy dom «Vil'jams», 2003. – 864 s.