

КОНЦЕПТУАЛЬНЫЕ ГРАФЫ И СЕМАНТИЧЕСКИЕ СЕТИ В СИСТЕМАХ ОБРАБОТКИ ЕСТЕСТВЕННО-ЯЗЫКОВОЙ ИНФОРМАЦИИ

Abstract: In this work, the approach to adaptation of known descriptions of conceptual graphs with reference to visualization of the formalized representation of naturally-language statements for inflectional languages is considered.

Key words: semantic networks, conceptual graphs, logic of predicates of the first order.

Анотація: У роботі розглянуто підхід до адаптації відомих описів концептуальних графів стосовно візуалізації формалізованого подання природничо-мовних висловлювань для флективних мов.

Ключові слова: семантичні мережі, концептуальні графи, логіка предикатів першого порядку.

Аннотация: В работе рассмотрен подход к адаптации известных описаний концептуальных графов применительно к визуализации формализованного представления естественно-языковых высказываний для флективных языков.

Ключевые слова: семантические сети, концептуальные графы, логика предикатов первого порядка.

1. Введение

Проблема поиска, представления, понимания и компьютерной обработки естественно-языковых объектов (ЕЯО) является одной из наиболее сложных проблем в свете решения соответствующих задач искусственного интеллекта, в частности, построения эффективного человеко-машинного общения на языке, близком к естественному.

Сложность этой проблемы состоит в том, что ЕЯО плохо формализуются и вследствие этого плохо поддаются автоматизации. В настоящее время существуют только некоторые частные подходы к формализации и автоматизации при решении этой проблемы. Как правило, эти подходы основываются на использовании языка формальной математической логики (классической и неклассической). Наиболее часто используется логика предикатов первого порядка с применением правила резолюций в качестве основного правила вывода и другими вспомогательными правилами. Особенно активно в последнее десятилетие развивается подход к представлению и анализу естественно-языковых объектов (ЕЯО) с использованием темпоральной и экстенциональной логик. В этих логиках правило резолюций не работает, и тут используются теоретико-автоматные методы исследования, методы семантического табло либо равносильные им.

2. Постановка задачи

В данной работе дан краткий обзор метода представления и обработки ЕЯО с использованием языка предикатов первого порядка. Данный формализм охватывает проблему анализа текстов в сочетании с методом резолюций.

Уместно уточнить, что проблему компьютерной обработки ЕЯО мы подразделяем на три подпроблемы:

- а) представления ЕЯО;
- б) получения следствий, проверки выполнимости, противоречивости или непротиворечивости заданного множества формул на основе выбранного представления;
- в) получения цепочки формального доказательства выводимости следствий, выполнимости (невыполнимости) заданного множества следствий.

Анализ перечисленных подпроблем показывает, что использование формальных методов проверки непротиворечивости или выполнимости множества формул в системах гильбертовского типа (т.е. систем, в которых доказательства получаются формальным образом из аксиом путем применения правил вывода) для подпроблемы в) не имеет более-менее удовлетворительного решения. Дело в том, что системы гильбертовского типа не являются структурированными, т. е. для сбора необходимой информации по одному единственному объекту необходимо просматривать все множество логических формул, которые находятся в системе (обычно такой системой является реляционная база данных). К сожалению, этим недостатком страдают все формальные логические системы гильбертовского типа. В целях устранения этого недостатка было предложено графическое представление формул и их аргументов, которое служит глобализации и структурированию информации. Основой графического представления являются концептуальные графы (КГ) и более сложные структуры – семантические сети (СС). Такое представление позволяет визуализировать модель естественно-языковой картины мира и, в случае необходимости, весь процесс доказательства или процесс вывода.

3. Неформальные определения и примеры

Введем формальные определения упомянутых выше понятий КГ и СС, а также приведем примеры для иллюстраций этих понятий.

Определение 1. Концептуальным графом называется размеченный двудольный ориентированный граф $G = (V_1 \cup V_2, E)$, где $V_1 \cap V_2 = \emptyset$, вершины из V_1 размечены именами предикатов, а вершины из V_2 – именами аргументов; E – множество дуг (ориентированных ребер). Дуги графа соединяют вершины, помеченные именами предикатов, с вершинами, помеченными именами аргументов. Вершины из множества V_1 называются узлами-предикатами, вершины из V_2 – узлами-концептами, а сами предикаты – концептуальными предикатами.

Из приведенного определения следует, что концептуальный граф должен удовлетворять таким условиям:

- 1) количество дуг, связывающих узлы-предикаты с узлами-концептами, равно арности предиката (т.е. числу его аргументов);
- 2) все узлы-предикаты, помеченные символом одного и того же концептуального предиката, имеют одинаковую арность;
- 3) все дуги, связывающие узлы-предикаты и узлы-концепты в КГ, упорядочены от 1 до n , где n – арность предиката;
- 4) графически узлы-предикаты обозначаются эллипсами, а узлы-концепты – прямоугольниками.

Если используемый логический язык типизирован, т.е. его объектам приписаны некоторые типы, то с каждым концептуальным предикатным символом связывается кортеж $\langle a, b, \dots, c \rangle$, называемый сигнатурой этого предиката. Или, другими словами, сигнатура предиката описывает семантическую модель правильного употребления аргументов в синтаксической структуре

некоторого ЕЯ-предложения с максимальной выразительной мощностью. Под максимальной выразительной мощностью понимается соответствующий кортеж без пустых элементов.

КГ не случайно выбраны в качестве структур данных для представления предикатов и их аргументов. Эти структуры прошли длительную апробацию в системах баз данных и зарекомендовали себя с наилучшей стороны. Это проявляется, прежде всего, в эффективности выполнения операций на таких структурах данных. Операции на множестве КГ выполняются с целью эффективного построения более сложной структуры – семантической сети.

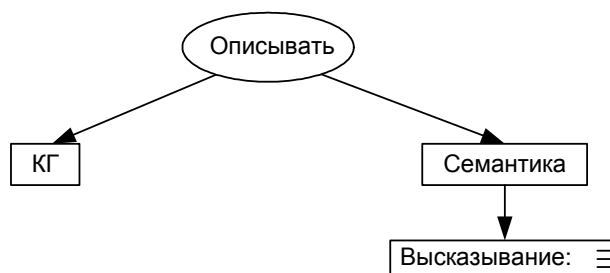
Определение 2. Семантической сетью называется объединение заданного множества концептуальных графов вместе с описанием их взаимосвязей и погружением в контекст области рассуждений.

Для подкрепления этих полуформальных определений КГ и СС рассмотрим примеры, иллюстрирующие эти понятия.

Пример 1. Рассмотрим предложение

“КГ описывает семантику некоторого высказывания”. (1)

Соответствующий этому предложению концептуальный граф представлен ниже.



Заметим, что КГ может иметь различные представления, которые включают некоторые уточнения, характеристики контекста или семантики.

В то время, как КГ представляют одну логическую формулу, СС представляет совокупность КГ, погруженных во взаимосвязи и общий контекст рассуждений. СС получаются из КГ по правилам конъюнкции и упрощения. Эти формальные правила сводятся к следующему: пусть дано два КГ G_1 и G_2 , тогда КГ G строится из этих графов путем применения правил конъюнкции и упрощения.

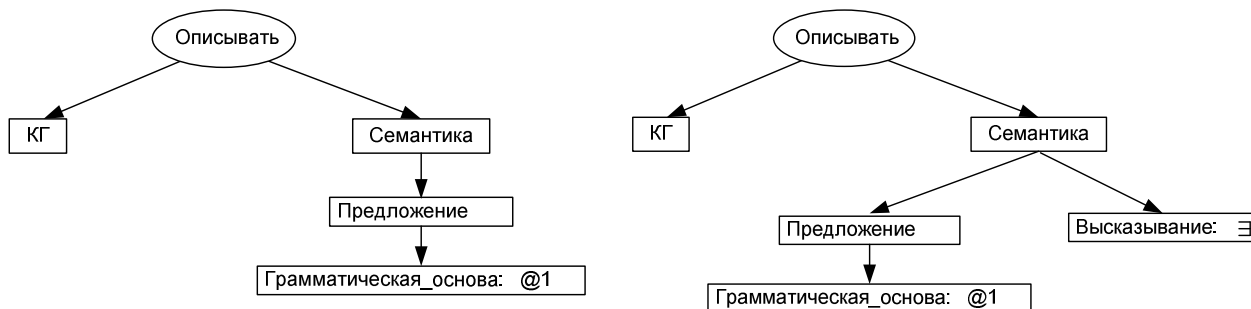
Правило конъюнкции (ПК): Если узел-концепт c_1 в G_1 идентичен узлу-концепту c_2 в G_2 , то G получается удалением c_2 и соединением с c_1 всех связывающих узлов, которые были связаны с c_2 в G_2 .

Правило упрощения (ПУ): Если КГ G после соединения содержит два идентичных (соединенных с одними и теми же узлами-концептами) связывающих узла, то можно удалить один из них вместе со связанными с ним дугами.

Пример 2. В качестве первого графа возьмем КГ из примера 1, а второй КГ соответствует предложению

“КГ описывает семантику предложения с одной грамматической основой”. (2)

В результате применения ПУ к концептуальным графам высказываний (1) и (2) получим семантическую сеть для указанных концептуальных графов, показанную ниже. Иногда такую СС называют объединенным КГ.



Далее основное внимание будет сосредоточено на КГ и методах их представления, типизации и обработки.

4. Способы представления КГ и их характеристика

Представление семантической модели ЕЯО в зависимости от типов решаемых задач (и соответствующих КГ) может иметь различную степень детализации. При этом различают две задачи формально-логического представления ЕЯО в процессе компьютерного анализа и интерпретации.

Первая из них относится к внутриязыковой обработке, при которой результат семантического анализа представляется наиболее полными логическими выражениями в соответствующем логическом базисе. Их формирование осуществляется параллельно с процедурой снятия лексической неоднозначности, которая предполагает детализацию КГ и эксплицитное представление соответствующих контекстных зависимостей.

Вторая задача относится к внеязыковой (или постязыковой, машинной) обработке, к этапу создания базы знаний (БЗ) предметной области (ПдО), а точнее – к созданию базы правил логического вывода. Такая база должна содержать короткие правила (занимать мало места в памяти), позволяющие реализовывать эффективный вывод (с точки зрения быстродействия) [3].

Из анализа указанных задач следует, что для первой задачи нужно использовать детализированные КГ, а для второй – обобщенные КГ и соответствующие СС.

Ниже будут рассмотрены некоторые определения и правила построения детализированных КГ, базирующиеся на концептуальных графах Дж. Соуы [1, 2]. Также будут рассмотрены основные проблемы, связанные с автоматическим построением детализированных КГ (описывающих входной ЕЯО), и предложены пути их преодоления.

Для наглядности описания детализированных концептуальных графов (далее концептуальных графов) будем использовать естественно-языковые высказывания (ЕЯВ) из ПдО “Интеллектуальные информационные системы (ИИС)” и определение семантической сети, взятое из [3]. Формирование структуры КГ, соответствующей указанному определению, будем проводить в три этапа: от простых трех высказываний (составляющих ядро определения семантической сети) –

на первом этапе, к их расширенным аналогам – на втором этапе и их объединению (полному определению) – на третьем.

Рассмотрим первое высказывание первого этапа (из упомянутого определения семантической сети) и соответствующий ему КГ (рис.1).

I.1. “Сети представляются в виде графа”.

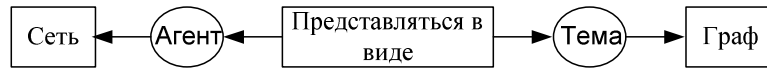


Рис. 1. Концептуальный граф для высказывания I.1

Существует и линейная форма записи КГ. При этом прямоугольники заменяются квадратными скобками, а эллипсы – круглыми.

[Сеть] ← (Агент) ← [Представляться в виде] → (Тема) → [Граф].

Обе формы записи: графическая и линейная, – представляют один и тот же абстрактный граф. Существует алгоритм, с помощью которого можно каждому КГ сопоставить формулу логики предикатов первого порядка. С помощью этого алгоритма КГ на рис. 1 преобразуется в нижеследующую формулу логики предикатов. Причем, если никакой другой квантор не определен в поле концепта, то квантором по умолчанию является квантор существования \exists .

$(\exists x: \text{Сеть})(\exists y: \text{Граф})(\exists z: \text{Представлять_в_виде})(\text{Агент}(z,x) \wedge \text{Тема}(z,y))$.

Неформально КГ – некоторая структура концептов и концептуальных отношений, где каждая дуга связывает некоторое концептуальное отношение R с некоторым концептом C .

Такое представление детализированных КГ несколько отличается от введенного выше описания. Объединение узлов концептов-предикатов и узлов концептов-аргументов в узлы-концепты и введение другого вида узлов – узлов-концептуальных отношений (рис. 1: агент, тема) является одним из главных моментов сути детализации. При этом узел концепта-предиката назовём главным узлом-концептом, а узлы концептов-аргументов – просто узлами-концептами. Такое преобразование, по сути, детализирует семантические отношения в синтаксической структуре предложения и представляет их в явном виде.

Компоненты структуры КГ определяются нижеследующими определениями:

- Каждая дуга E в G связывает концептуальное отношение R с концептом C . Дуга E принадлежит отношению R и не принадлежит концепту C .
- Концептуальный граф G может иметь концепты, которые не связаны ни с каким концептуальным отношением, но любая дуга, принадлежащая некоторому концептуальному отношению в G , должна быть присоединена только к одному концепту.
- Подграфами концептуального графа G могут быть:
 - а) пустой граф, который не имеет ни концептов, ни концептуальных отношений, ни дуг;
 - б) одноэлементный граф, который состоит из одного концепта;
 - в) граф-звезда, который состоит из одного концептуального отношения и концептов, присоединенных к его дугам.

Рассмотрим следующие два высказывания первого этапа и концептуальные графы для них запишем в линейной форме (рис. 2).

I.2. “Вершины соответствуют объектам”.

[Вершина] ← (Агент) ← [Соответствовать] → (Тема) → [Объект].

I.3. “Дуги соответствуют отношениям”.

[Дуга] ← (Агент) ← [Соответствовать] → (Тема) → [Отношение].

Рис. 2. Высказывания I.2 и I.3 и их концептуальные графы

Второй этап

Каждый концепт C имеет тип t и референт r . Слева записывается концепт-тип, а через двоеточие справа записывается подтип (элемент, терм, конкретное имя), который присутствует в высказывании. А теперь рассмотрим концептуальный граф для высказывания “Сеть высказывания I.1 представляется в виде концептуального графа”. Для него концепт [Сеть] примет вид [Сеть: Высказывание_I.1]. Такое представление предполагает наличие онтологии ПДО “ИИС” (и базы

знаний, содержащей список высказываний, в том числе I.1), фрагмент которой представлен на рис. 3.

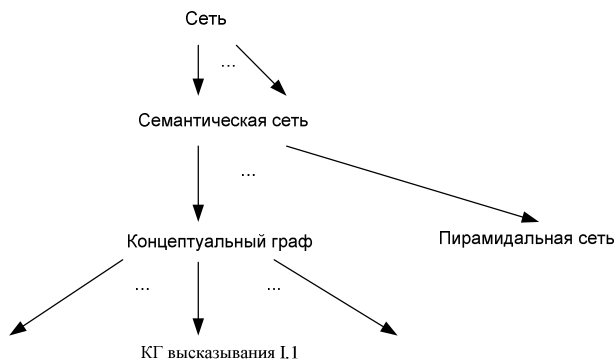


Рис. 3. Фрагмент онтологии концепта “Сеть” из ПДО “ИИС”

Каждое концептуальное отношение R имеет реляционный тип t и целое неотрицательное число n , названное арностью.

Число дуг, которые принадлежат R , равно арности n , при этом сами дуги нумеруются от 1 до n .

Для каждого n -арного R существует последовательность, состоящая из n концептных типов $\langle c_1, c_2, \dots, c_n \rangle$, названная сигнатурой R . Сигнатура $R \langle c_1, c_2, \dots, c_n \rangle$ является частным случаем сигнатуры концепта-предиката $\langle a, b, \dots, c \rangle$. Сигнатура отношения в общем виде представляет ограничение на типы

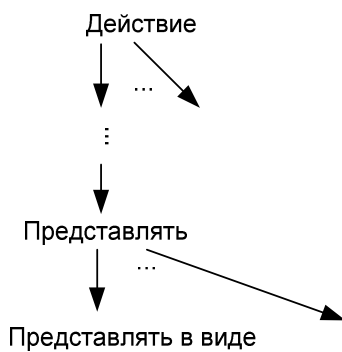


Рис. 4. Фрагмент онтологии для концепта-типа “Действие”

концептов, которые могут быть связаны с его дугами. Например, для отношения “Агент” сигнатура – $\langle \text{Действие}, \text{Сущность} \rangle$, которая указывает, что тип концепта, связанного с его правой дугой, должен быть “Действие” или некоторым его подтипом в онтологической иерархии. Для высказывания I.1 используемое отношение “Агент” предполагает следующую онтологическую иерархию (рис. 4).

Концепты-типы и концептуальные отношения имеют иерархические структуры, которые представляют собой

частично упорядоченные множества T и R соответственно, чьи элементы обозначают типы концептов и отношений. Каждое обозначение концепта в T и отношения в R определено как примитив или как частичное упорядочение по T и R отношениям подтипа с символом \leq – для подтипа, $<$ – для собственного подтипа, \geq – для супертипа и $>$ – для собственного супертипа.

Референт концепта определен квантором и указателем. Квантором может быть или квантор существования или квантор всеобщности или обозначения: количества – (@1) и набора {1,2,3}.

Квантор существования представлен или символом \exists или отсутствием любого другого символа квантора или выражения.

Квантор всеобщности представлен или символом \forall или выражением в расширенном синтаксисе, которое может быть оттранслировано в концептуальные графы, содержащие только кванторы существования.

Указатель в поле референта может быть трех видов.

Литерал – синтаксическое представление формы референта. Литералы также бывают трех видов – числа, строки символов и закодированные литералы, которые определены парой, состоящей из идентификатора и строки.

Локатор – символ, который определяет, как может быть найден референт. Существуют два вида локаторов, отличающихся способом определения референта: индивидуальный маркер (определяет уникальный концепт в каталоге экземпляров БЗ) и индексированный маркер (символ, определяющий референт в соответствии с некоторыми соглашениями, которые являются независимыми от текущей БЗ), начинающийся с символа #.

Дескриптор – концептуальный граф, который описывает референт.

Референт концепта определяет сущность или множество сущностей, к которым относится концепт.

Вместо обращения к объекту по маркеру или по имени можно использовать обобщенный маркер “*” для определения неспецифицированных экземпляров. Эта информация часто опускается в метках понятий: узел, заданный только меткой типа [Вершина], эквивалентен узлу [Вершина: *]. Дополнительно к обобщенному маркеру КГ позволяют использовать именованные переменные. Они представляются звездочкой, за которой следует имя переменной, например, [Вершина: *x]. Это позволяет указывать в разных концептуальных графах один и тот же обобщенный концепт-тип.

В КГ под контекстом понимается концепт с вложенным КГ, который описывает референт. В свою очередь, вложенный КГ может содержать концепты с вложенными концептуальными графами, т.е. концепты могут быть вложенными.

Кореферентным множеством C в концептуальном графе G называется множество концептов, отобранных из G или из графов, вложенных в контексты G .

В графической системе обозначений элементы кореферентного множества могут быть соединены пунктирными линиями, которые называются кореферентными связями.

В линейной системе обозначений два концепта могут быть связаны пунктирной линией только тогда, когда они записаны рядом друг с другом в одной и той же строке.

В ЕЯВ кореферентная связь обычно выражается некоторой формой глагола *быть*. Например, КГ $\boxed{\text{КГ}}$ — — — — $\boxed{\text{ЕЯВ}}$ может быть прочитан так: “Некоторый концептуальный граф есть естественно-языковое высказывание”. В исчислении предикатов кореферентная связь соответствует равенству $(\exists x: \text{КГ})(\exists y: \text{ЕЯВ}) x = y$.

В теоретическом базисе концептуальных графов определено *понятие базы знаний*, под которым подразумевается контекст типа KnowledgeBase, десигнатором его является некоторый полифункциональный КГ, состоящий из четырех компонентов:

- иерархия типов. Контекст типа TypeHierarchy, десигнатором которого является КГ T, который определяет частичное упорядочение обозначений типа и одноместных лямбда-выражений для каждого определенного обозначения типа;
- иерархия отношений. Контекст типа RelationHierarchy, десигнатором которого является КГ R, который определяет частичное упорядочение отношений, а также арности для каждого отношения;
- каталог используемых индивидов. Контекст типа CatalogofIndividuals, десигнатором которого является КГ C, содержащий точно один концепт для каждого индивидуального маркера КГ C и может быть использован в БЗ. Десигнатор может содержать другие концепты и отношения, что описывают индивиды;
- верхний уровень контекста (контекст типа Assertion), десигнатором которого является КГ A.

Содержание БЗ должно удовлетворять следующим ограничениям:

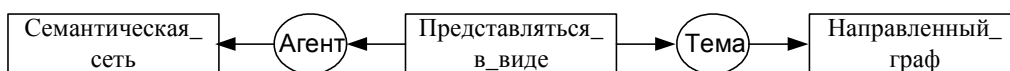
- обозначения типов в любом концептуальном графе должны быть специфицированы в T ;
- обозначения отношений в любом концептуальном графе должны быть специфицированы в R ;
- индивидуальные маркеры в любом концептуальном графе должны быть специфицированы в C .

Обозначения типов KnowledgeBase, TypeHierarchy, RelationHierarchy, CatalogofIndividuals и Assertion являются метауровневыми обозначениями и не специфицируются в T .

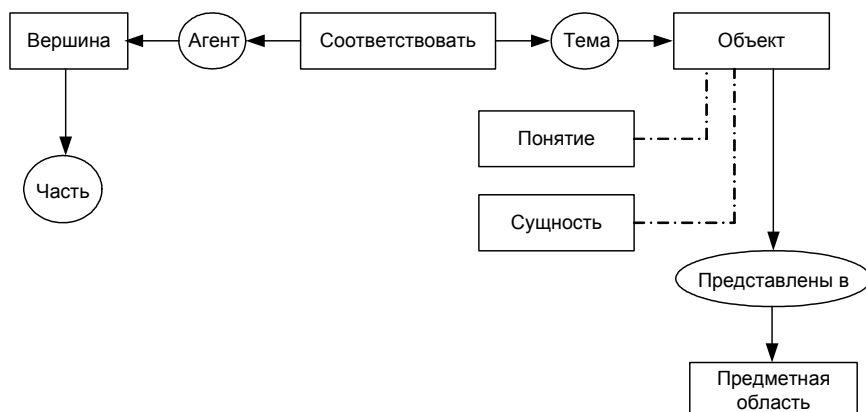
Приведенное определение соотносит БЗ с единственным концептом, который также может быть вложен в другую базу знаний. Каталог индивидов перечисляет все именованные сущности в выбранной для исследования предметной области.

Для дальнейшего рассмотрения основ построения концептуальных графов на рис. 5 представлены расширенные высказывания второго этапа и соответствующие им концептуальные графы.

II.1. “Семантические сети представляются в виде направленного графа”.



II.2. “Вершины направленного графа соответствуют объектам (понятиям, сущностям) предметной области”.



II.3. “Дуги направленного графа соответствуют отношениям (связям) между объектами”.

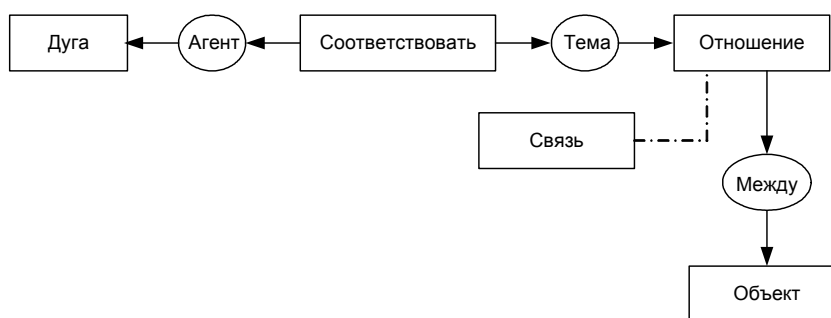


Рис. 5. Расширенные высказывания второго этапа и их КГ

В расширенных ЕЯВ второго этапа введенные (на первом этапе) концепты-типы конкретизируются, для некоторых из них добавляются новые концептуальные отношения и вводятся понятия-синонимы. Указанные расширенные ЕЯВ представляют собой три составляющие (три грамматические основы) одного из принятых формальных определений семантической сети.

На третьем этапе, объединив соответствующие КГ, мы получим результирующий КГ (или семантическую сеть) для указанного определения.

Несмотря на огромную популярность концептуальных графов, в известной методологии их построения [1, 2] имеется ряд задач, при решении которых у разработчика конкретной знание-ориентированной компьютерной системы возникают серьезные трудности, разрешение которых порой оказывается проблематичным. Например, в случае компьютерной обработки естественно-языковых объектов (ЕЯО):

- не определены принципы построения иерархии концептов-типов и концептуальных отношений, а также механизмы их обработки;
- отсутствует механизм функционального анализа ЕЯВ (определения ролевых функций членов предложения);
- не определен механизм выделения терминов из ЕЯО, описывающих некоторую ПдО, причем термины могут состоять из нескольких словоформ различных частей речи (например, “семантическая сеть”, “направленный граф” и др.);

- не определен механизм разрешения индексированных меток (или проблемы анафорических связей);
- не определен механизм фильтрации или отбора из терминов рассматриваемой ПдО концептов и концептов-индивидов;
- не определены механизмы системной интеграции лингвистического и предметного аспектов анализа ЕЯО;
- не определены механизмы квантификации переменных;
- введенное понятие кореферентного множества сформулировано неоднозначно для понимания;
- не определены механизмы формирования шаблонов сигнатур для концептуальных отношений.

Решение проблем, возникающих в связи с устранением описанных выше недостатков, возможно на основе принципов и механизмов построения языково-онтологической информационной системы (ЯОИС). Подробно ЯОИС рассматривается в ряде работ [4, 5 и др.]. Применительно к проблеме построения концептуальных графов в ЯОИС задействованы компоненты: “Языково-онтологическая картина мира (ЯОКМ)”, “Онтология ПдО” и “Грамматический процессор”. ЯОКМ представляет собой общелингвистическую онтологию, компонентами которой являются онтологии лексем таких частей речи, как существительное, глагол, прилагательное и наречие. Онтология ПдО представляет собой онтологическую структуру концептов-типов ПдО, специфичных для нее концептуальных отношений и лексических средств. Системная интеграция указанных онтологий (лингвистической и предметной области) совместно с базой фактов и правил вывода для рассматриваемой ПдО содержит в себе в некотором виде концепт KnowledgeBase, а грамматический процессор формирует структурные единицы соответствующих иерархий.

Концепты различаются по разновидностям, тесно связанным с грамматическими характеристиками ЕЯВ. При построении алгоритма интерпретации ЕЯВ и его КГ важная роль отводится глаголам-сказуемым, поскольку они определяют отношения между подлежащим, дополнением и другими компонентами предложения. Помимо концептов, необходимо определить концептуальные отношения, которые будут использованы в КГ. В общем случае для связи компонент в предложении верхний уровень иерархии отношений можно определить следующим образом:

- Слово *Агент* служит для обозначения отношения, связывающего концепт *Действие* с концептом *Сущность*. Здесь следует различать концепт *Сущность* как одушевленный объект и неодушевленный. Следовательно, применяются разные интерпретации отношения *Агент*.
- Слово *Носитель_состояния* служит для обозначения отношения, связывающего концепт *Состояние* с концептом *Одушевленный*, отражающее связь между психическим состоянием и его носителем.
- Слово *Объект* служит для обозначения отношения, связывающего концепт *Событие* или *Состояние* с концептом *Сущность*.

- Слово *Инструмент* служит для обозначения отношения, связывающего концепт *Действие* с концептом *Сущность*, и определяет инструмент, используемый для выполнения действия.

- Слово *Часть* служит для обозначения отношения, связывающего концепты “Часть-целое”.

- Слово *Тема* служит для обозначения отношения, связывающего концепт *Событие* или *Состояние* с концептом *Сущность* некоторой абстрактной формы.

- Слово *Время* маркирует время действия глагола-сказуемого в предложении.

- Слово *Цель* связывает концепт с другим концептом, который представляет его цель.

- Слово *Пациент* связывает концепт *Действие* с концептом *Одушевленный*, на кого направлено действие.

- Слово *Способ* связывает концепт *Действие* с концептом, обозначающим способ выполнения действия (в предложении обычно выражается наречием).

- Слова, обозначающие предлоги (*в, на и др.*) маркируют отношения семантики падежных форм существительных.

- И другие.

Предположим, что указанная обработка ЕЯВ второго этапа уже выполнена. Тогда, применив операцию объединения к концептуальным графам (рис. 5), получим синтезированную СС для третьего этапа (рис. 6).

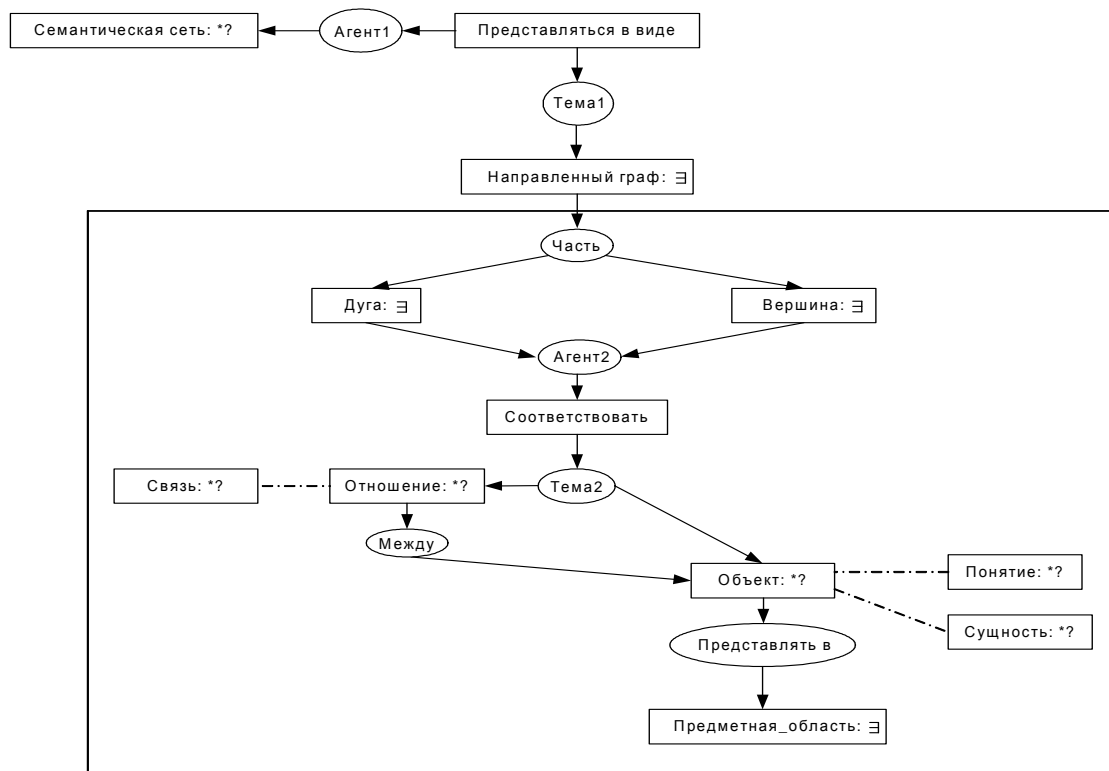


Рис. 6. СС ЕЯВ третьего этапа

Она может быть прочитана следующим образом.

Семантические сети представляются в виде направленного графа, вершины которого соответствуют объектам (понятиям, сущностям) предметной области, а дуги – отношениям (связям) между объектами.

Для формирования шаблонов сигнатур концептуальных отношений требуется разработать алгоритм их извлечения из интегрированной онтологии “ЯОКМ∧ Онтология_ПдО” (где они представлены в неявном виде) и эксплицитного представления.

Концептуальные графы могут включать концепт-тип [Суждение], объектом ссылки которого является множество концептуальных графов. Этот концепт изображается прямоугольником, содержащим указанное множество концептуальных графов.

КГ позволяют легко представить такие конъюнктивные понятия, как “Граф содержит вершины и дуги”. Сложнее обстоит вопрос с операциями отрицания и дизъюнкции.

Операция отрицания “Не” вводится как операция, аргументом которой является концепт-тип [Суждение]. При этом высказывание становится ложным. Так, высказывание “Не существует графов без вершин и дуг” может быть представлено следующим образом.

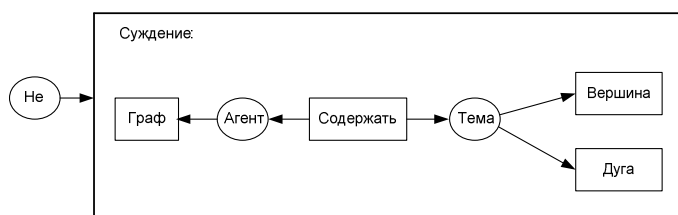


Рис. 7. КГ с операцией отрицания

Используя отрицание и конъюнкцию, можно формировать КГ для представления дизъюнктивных утверждений в соответствии с правилами логики. Предполагается, что в КГ родовые понятия связаны

квантором существования. Используя отрицание и квантор существования, можно представить квантор всеобщности. Например, КГ на рис. 7 можно рассматривать как представление логического выражения

$$\forall X1 \forall X2 \forall X3 \forall X4 \left(\neg \left(\text{Граф}(X1) \wedge \text{Вершина}(X2) \wedge \text{Дуга}(X3) \wedge \text{Содержать}(X4) \wedge \left(\text{Агент}(X4, X1) \wedge \text{Тема}(X4, (X2 \wedge X3)) \right) \right) \right)$$

Концептуальные графы по своей выразительной мощности эквивалентны исчислению предикатов первого порядка. Существует прямое соответствие представления КГ системе обозначений исчисления предикатов. В [2] предложен алгоритм представления некоторого концептуального графа G формулой исчисления предикатов. Суть его сводится к выполнению таких шагов:

1. Каждое из n родовых концептов G связать с отдельной переменной x_1, x_2, \dots, x_n .
2. Каждое конкретное понятие в G связать с отдельной константой. Эта константа может быть просто именем или маркером, используемым для указания объекта ссылки для данного понятия.
3. Представить каждый узел концепта унарным предикатом, имя которого соответствует типу этого узла, аргументом которого является переменная или константа данного узла.
4. Представить каждое n -арное концептуальное отношение на графе G n -арным предикатом с тем же именем. Каждый аргумент предиката является переменной или константой, соответствующей узлу концепта, связанного с этим отношением.
5. Сформировать тело выражения исчисления предикатов в виде конъюнкции атомарных выражений, построенных в п.п. 3.4. Все переменные в выражении считаются связанными квантором существования.

Для примера приведем конечные выражения исчисления предикатов для КГ ЕЯВ второго этапа (рис. 5), полученные в результате применения описанного выше алгоритма. При этом:

– родовые концепты КГ для ЕЯВ II.1 связаны с переменными: X1 – “Семантическая_сеть”, X2 – “Направленный_граф”, X11 – “Представляться_в_виде”;

– родовые концепты КГ для ЕЯВ II.2 связаны с переменными: X3 – “Вершина”, X4 – “Объект”, X5 – “Понятие”, X6 – “Сущность”, X7 – “Предметная_область”, X12 – “Соответствовать”;

– родовые концепты КГ для ЕЯВ II.3 связаны с переменными: X8 – “Дуга”, X9 – “Отношение”, X10 – “Связь”, X12 – “Соответствовать”.

II.1.

$$\exists X1 \exists X2 \exists X11 \left(\left(\left(\text{Семантическая_сеть}(X1) \wedge \text{Направленный_граф}(X2) \wedge \right. \right. \right. \\ \left. \left. \left. \wedge \text{Представляться_в_виде}(X11) \wedge \text{Агент1}(X11, X1) \wedge \text{Тема1}(X11, X2) \right) \right) \right).$$

II.2.

$$\exists X3 \exists X4 \exists X5 \exists X6 \exists X7 \exists X12 \left(\left(\left(\text{Вершина}(X3) \wedge \text{Объект}(X4) \wedge \text{Понятие}(X5) \wedge \text{Сущность}(X6) \wedge \right. \right. \right. \\ \left. \left. \left. \wedge \text{Предметная_область}(X7) \wedge \text{Соответствовать}(X12) \wedge \right. \right. \right. \\ \left. \left. \left. \text{Агент2}(X12, X3) \wedge \text{Тема2}(X12, X4) \wedge \text{Часть}(X3, X2) \wedge \right. \right. \right. \\ \left. \left. \left. \wedge \text{Представлять_в}(X4, X7) \wedge \text{Быть}(X4, X5) \wedge \text{Быть}(X4, X6) \right) \right) \right).$$

II.3.

$$\exists X8 \exists X9 \exists X10 \exists X12 \left(\left(\left(\text{Дуга}(X8) \wedge \text{Отношение}(X9) \wedge \text{Связь}(X10) \wedge \text{Соответствовать}(X12) \wedge \right. \right. \right. \\ \left. \left. \left. \wedge \text{Агент2}(X12, X8) \wedge \text{Тема2}(X12, X9) \wedge \text{Быть}(X9, X10) \right) \right) \right).$$

5. Выводы

В работе рассмотрен подход к адаптации известных описаний КГ декларативного характера применительно к визуализации формализованного представления ЕЯВ для флективных (например, русского или украинского) языков. Основным назначением построения КГ является создание компьютерных систем, которые предоставляли бы возможность явного представления знаний, извлеченных из ЕЯО, и позволили бы их анализировать. Показана недостаточность выразительной мощности известных представлений КГ при создании прикладных систем, а также предложены некоторые пути по ее преодолению. Суть предложений сводится к применению онтологического подхода к проектированию КГ в системе описательных средств компьютерной системы, в частности, использования общелингвистической и предметной онтологий.

СПИСОК ЛИТЕРАТУРЫ

1. Sowa, John F. Knowledge Representation: Logical, Philosophical, and Computational Foundations. – Brooks Cole Publishing Co., Pacific Grove, CA, 2000. – 594 p.
2. Sowa J.F. Conceptual Structures: Information Processing in Mind and Machine. Reading, MA: Addison – Wesley, 1984.
3. Палагин А.В., Яковлев Ю.С. Системная интеграция средств компьютерной техники. – Винница: «УНІВЕРСУМ-Вінниця», 2005. – 680 с.
4. Палагін О.В., Петренко М.Г. Розбудова абстрактної моделі мовно-онтологічної інформаційної системи // Математичні машини і системи. – 2007. – № 1. – С. 42 – 50.
5. Палагин А.В., Петренко Н.Г. К вопросу системно-онтологической интеграции знаний предметной области // Математичні машини і системи. – 2007. – № 3, 4. – С. 63 – 75.

Стаття надійшла до редакції 20.01.2009