



УДК 004.822:514

УНИФИЦИРОВАННОЕ ПРЕДСТАВЛЕНИЕ И ИНТЕГРАЦИЯ ЗНАНИЙ

Ивашенко В.П.

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

ivashenko@bsuir.by

Рассматриваются составляющие и применение средств технологии компонентного проектирования баз знаний в виде однородных семантических сети с теоретико-множественной семантикой для решения задач отладки и интеграции баз знаний.

Ключевые слова: база знаний, онтология, семантическая сеть, интеграция.

ограниченность языков и моделей представления знаний

ВВЕДЕНИЕ

Рассматриваются задачи интеграции знаний в рамках задачи разработки базы знаний для интеллектуальной системы [Ивашенко, 2009a].

Интеграция знаний заключается в погружении новых знаний в состав знаний, уже известных интеллектуальной системе, и является необходимым этапом понимания информации, которая либо поступает в интеллектуальную систему извне, либо генерируется (порождается) в процессе обработки информации. Процесс понимания в интеллектуальной системе заключается не только и не столько в помещении добавляемых новых знаний в память, но – в выявлении синонимов, связывании и отождествлении поступающих фрагментов с уже известными. Таким образом, интеграция знаний является основой функционирования любой интеллектуальной системы и во многом определяет качество интеллектуальной системы, являясь последним и основным этапом понимания поступающих (приобретаемых) знаний.

Разработка баз знаний наталкивается на следующие трудности:

- отсутствие развитых технологий разработки или их ограничения:
 - ограниченные возможности верификации и отладки баз знаний (отсутствие средств верификации количества синонимичных знаков, неполное выявление противоречий);
 - плохая отчуждаемость и переносимость базы знаний в силу привязки к инструментарию или конкретной оболочке (CLIPS) [CLIPS, 1991],

- немногочисленность инженеров баз знаний (из-за высоких стартовых требований к разработчику) – от разработчика требуется владение специальными знаниями по моделям и языкам представления знаний):

- необходимость выбора среди нескольких моделей знаний, между которыми нет однозначного предпочтения [Martin Ph., 2002];
- неоднородность моделей представления знаний, приводящая к тому, что в базе появляются синонимичные структуры разных типов, требующие от пользователя их согласования;
- сложность языков представления знаний
- не полностью решён вопрос интеграции баз знаний:
 - ограничения на расширение базы знаний;
 - ограниченные иерархической таксономической структурой онтологии возможности интеграции баз знаний или отсутствие таких возможностей, необходимость выбора средств интеграции пользователем [Gangemi et al., 1996];
 - отсутствие общих стандартов совместимости разработанных фрагментов баз знаний;
 - ограниченность средств поиска и каталогизации разработанных фрагментов баз знаний;
 - ограниченность или отсутствие простых средств взаимодействия с внешней средой.

Для разработки баз знаний используются:

- языки представления и обработки баз знаний;
- средства создания и отладки баз знаний;
- средства интеграции баз знаний.

К языкам представления знаний в разных системах на сегодняшний день относятся: Conceptual Graph [Sowa et al., 2008], Frame-logic [Michael Kifer et al., 1995], Knowledge Interchange Format [Genesereth et al., 1992], Integrated Definition for Ontology Description Capture Method (IDEF5) [IDEF5, 1994], Common Algebraic Specification Language [CoFI:CASL-Summary, 2004], Concept maps/UML [Novak et al., 2008] [ISO24707], RDF/RDFS [W3C:RDFS, 2004], DARPA Agent Markup Language [DAML, 2006], CycL [CycL, 2002], Ontology Inference Layer [W3C:DAML+OIL, 2001], RDF/OWL-Lite, RDF/OWL-DL, RDF/OWL Full [W3C:OWL, 2004], Topic Maps [ISO13250], XTM/LTM/CML/GML [XTM, 2001], Common Logic [ISO24707], Developing Ontology-Grounded Methods and Applications [Mustafa Jarrar et al., 2008], Formal English [Martin Ph., 2002], Gellish [Van Renssen, 2005], RDF/Rule Interchange Format [W3C:RIF, 2010], Open Biomedical Ontologies [Smith et al., 2007], RDF/OWL2 [W3C:OWL2, 2009], RDF/OWL2 EL, RDF/OWL2 RL и др.

Эти языки могут быть классифицированы по различным признакам:

- синтаксические признаки
 - URI-ориентированные
 - линейные языки
 - мультиплетные языки
 - триплетные языки
 - квинтиплетные языки
 - иерархически структурированные языки
 - ЛИСП-подобные языки
 - нелинейные языки
 - графовые языки
 - гиперграфовые языки
- семантические признаки
 - фреймовые языки
 - правила
 - семантическая сеть
 - семантический гиперграф
 - логические языки
 - дескрипционной логики
 - логики предикатов первого порядка
 - логики предикатов высших порядков
 - модальной логики
 - темпоральной логики (линейной, ...)
 - взаимодействующих процессов Хоара
 - объектно-ориентированные языки
 - процедурные языки
 - естественный язык

Кроме очевидных, дополнительно к недостаткам [Martin Ph., 2002] вышеперечисленных языков представления знаний можно отнести:

- отсутствие разделения понятий и терминов (исключения – DOGMA и Gellish),
- отсутствие поддержки монотонного расширения базы знаний (исключения – OWL, OWL2 и языки, построенные на основе классических логических моделей),
- отсутствие у некоторых из вышеперечисленных языков возможности семантического расширения языка (замкнутость языка).

Для поиска в базах знаний и онтологиях [Хорошевский, 2008] используются такие языки, как: RDQL, squish, SPARQL [W3C:SPARQL, 2008], KQML, DMX, Datalog, TSQL, ERROL, RuleML, RQL, OQL, TQL, VERSA, DQL и др.

Среди средств, которые могут рассматриваться в качестве основы для разработки баз знаний, можно выделить: оболочки экспертных систем (CLIPS (FuzzyCLIPS, DYNACLIPS, WxCLIPS) [CLIPS, 1991], SOAR, OPS83, RT-EXPERT, MIKE, BABYLON, WindExS, ES; ACQUARE, Easy Reasoner, ECLIPSE, EXSYS Professional, SIMER+MIR, AT ТЕХНОЛОГИЯ, CAKE v2.0) [Гаврилова и др., 2000]; инструментальные пакеты для разработки экспертных систем (G2, ART, KEE, Knowledge KRAFT); системы, ориентированные на обработку онтологий [Sowa et al., 2008] – Protégé, WebOnto, OntoEdit, WebODE, OilEd, OntoLingua.

Достоинствами приведённых средств являются: поддержка представления знаний различного вида различными моделями представления знаний в рамках одной системы; наличие средств визуального проектирования баз знаний; наличие средств верификации базы знаний, включая проверку на непротиворечивость; возможность монотонного расширения базы знаний, наличие средств интеграции баз знаний; наличие средств поддержки обмена данными с внешней средой, включая средства обмена данными в реальном времени.

Для преодоления трудностей семантической интеграции [Doan and Halevy, 2005], [Кудрявцев, 2008] (отображения онтологий (ontology mapping) и интеграции знаний (knowledge integration) в базах знаний используются следующие подходы:

- сравнение и выравнивание онтологий (ontology matching & alignment),
- интеграция онтологий (ontology merging),
- семантическое сравнение (semantic matching),
- семантическая унификация (semantic unification).

Для интеграции онтологий можно выделить методы:

- структурно-синтаксические:
 - анализ внутренней структуры
 - экстенционально-статистические
 - анализ внешней структуры (метаструктурный анализ)
 - терминологические (лексические)
- логико-семантические

К наиболее развитым подходам и методам интеграции можно отнести ONION (ONtology composition [Mitra, 2001]), формальный концептуальный анализ (FCA [Ganter, 1999], [Stumme, 2001]) и варианты его развития для нечётких (FOGA [IEEE, 2006]) и неопределённых множеств, методы использующие элементы семантического анализа – СТХМАТЧН [Bouquet, 2003] и S-match. Однако все эти методы, несмотря на использование некоторыми из них нечётких и неопределённых множеств, плохо приспособлены или не приспособлены к интеграции знаний в условиях наличия НЕ-факторов [Нариньяни, 2000]. Перечисленные методы, и методы, основанные на мерах близости (CUPID [Maldavan, 2001] и т.п.), не обеспечивают достаточной формальной строгости и непротиворечивости онтологий или баз знаний, получаемых в результате.

К существующим на настоящий момент средствам интеграции онтологий можно отнести: Optima, Prompt, Ontolingua, Chimaera [McGuinness et al., 2000], ONION [Gangemi et al., 1996], COMA++ [Aumueller et al., 2005] и др.

Таки образом, все эти средства и методы имеют ограничения и не преодолевают в полной мере вышеперечисленные трудности.

1. Модель унифицированного нелинейного представления знаний

Семантическая модель интеграции использует унифицированное представление знаний и обеспечивает интеграцию sc-моделей баз знаний и их фрагментов [Ивашенко, 2012]. Унифицированное представление знаний обеспечивается моделью унифицированного нелинейного представления знаний, которая является частным случаем такой модели представления знаний, как семантические сети и задаётся семейством совместимых sc-языков, использующих унифицированный способ семантического кодирования Semantic Computer code (SC-код) [Голенков и др, 2001] и поддерживающих представление знаний различного вида [Ивашенко, 2003], [Ивашенко, 2004], [Ивашенко, 2011a]. Особенности SC-кода являются: простой алфавит, содержащий узлы и дуги, простой синтаксис, базовая теоретико-множественная интерпретация. Семантика sc-языка задаётся на основе модели ситуативных множеств, являющейся развитием моделей L-нечётких множеств и неоднородных нечётких множеств, предложенных Дж. Гогеном и А. Кофманом [Кофман, 1982].

Модель ситуативных (событийных, нестационарных [Ивашенко, 2012]) множеств может быть задана следующей шестёркой компонентов:

$$\langle Universe, [0;1], Events, r, h, sM \rangle \quad (1)$$

где $Universe$ – универсальное множество объектов предметной области, $Events$ – множество

элементарных событий, $r \subseteq Events \times Events$ – отношение доступности (следования во времени) событий, $h \in (2^{Events})^{Universe}$ – функция, задающая множество событий существования каждого элемента универсального множества, sM – семейство пар множества событий существования ситуативного множества и функций (нечёткой) ситуативной принадлежности элементов универсального множества ситуативному множеству, отображающих элементы ситуативных множеств, множества событий и соответствующие им наборы степеней нечёткой принадлежности высших порядков на множество степеней нечёткой принадлежности $[0;1]$.

Для вычисления степеней нечётких принадлежностей введены параметризованные выражения для верхних и нижних треугольных норм и выразимых через них операций, задающих соответственно значения степеней принадлежности для операций пересечения, объединения и разности ситуативных множеств.

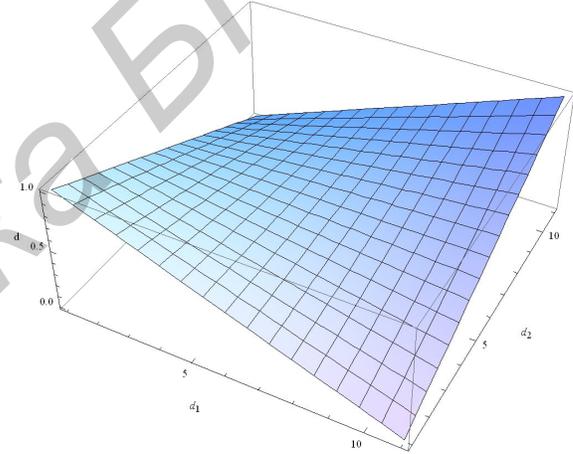


Рисунок 1– график функции нечёткой принадлежности элемента универсального нечёткого множества третьего порядка (d_1 и d_2 – аргументы, компоненты вектора нечётких степеней принадлежности, d – значение).

Выражение для верхней треугольной нормы: x и y – основные аргументы, значения принадлежностей исходных множеств, u – параметр, значение принадлежности универсального нечёткого множества (рис. 1).

$$tnormu(\langle x, y, u \rangle) = \max(\{x\} \cup \{y\} \cup \{u\}) + \min(\{x\} \cup \{y\} \cup \{u\}) - u \quad (2)$$

Нижняя s-норма.

$$snormn(\langle x, y, u \rangle) = x + y - tnormu(\langle x, y, u \rangle) \quad (3)$$

Верхняя s-норма.

$$snormu(\langle x, y, u \rangle) = snormn(\langle x + y + u - 1, u, (x + y) / 2 \rangle) \quad (4)$$

Верхняя разностная форма.

$$dformu(\langle x, y, u \rangle) = tnormu(\langle x, 1 - y, u \rangle) \quad (5)$$

Нижняя разностная форма.

$$dformn(\langle x, y, u \rangle) = x + 1 - u - tnormu(\langle x, y, u \rangle) \quad (6)$$

Нижняя треугольная норма.

$$tnormn(\langle x, y, u \rangle) = dformn(\langle x, 1 - y, u \rangle) \quad (7)$$

При переходе от ситуативных множеств низших порядков (нечёткости) к множествам высшим порядков используются выражения, позволяющие переычислять значения степени нечёткой принадлежности $degree(\langle b, V \rangle)$, исходя из модели равновероятных исходов в конечной темпоральной модели ветвящегося времени. Для этого используется базовое значение вероятности b , которое позволяет рассчитать степени принадлежности высших порядков по рекуррентному выражению для $f(V)$.

$$degree(\langle b, V \rangle) = b + (1 - 2 * b) * f(V) \quad (8)$$

$$f(V) = \min(\{1\} \cup \{V_1 * f(tail(V)) | \dim(V) > 0\}) + \max(\{0\} \cup \{(1 - V_1) * f(inverse(V)) | \dim(V) > 1\}) \quad (9)$$

$$inverse(V) \in \times_{i=2}^{\dim(V)} \{1 - V_i - \text{sgn}(i - 2) * (1 + 2 * V_i)\} \quad (10)$$

где $tail(V)$ – вектор всех компонентов вектора V , начиная со второго компонента, а $\dim(V)$ – размерность вектора V .

В соответствии с введённой моделью ситуативных множеств для описания динамических предметных областей, используются понятия нестационарной принадлежности и непринадлежности.

Модель унифицированного нелинейного представления знаний задаётся следующими компонентами.

$$\langle SClanguages, R_{sc}, F_{sc} \rangle \quad (11)$$

$$F_{sc} = A_{sc} \cup I_{sc} \cup N_{sc} \cup S_{sc} \cup E_{sc} \cup K_{sc}$$

$SClanguage$ – множество sc-языков, R_{sc} – отношения на множестве sc-языков (sc-подъязыка и трансляции), F_{sc} – функции sc-языков, A_{sc} – алфавитные функции sc-языков, I_{sc} – функции отношений инцидентности, N_{sc} – синтаксические предикаты sc-языков, S_{sc} – семантика sc-языков, E_{sc} – ключевые элементы sc-языков, K_{sc} – спецификация sc-языков и их ключевых элементов (отображение на множество онтологий sc-языков).

Semantic Code ориентирован на универсальность, т.е. на представление любых видов знаний. Semantic Code является средством унификации представляемых знаний. На множестве sc-языков (языков, представленных в

SC-коде) определены отношения sc-подъязыка и трансляции. Подъязык, являющийся пересечением выделенного семейства совместимых специализированных sc-языков, рассматривается как интегрированный sc-язык представления знаний (SCK). Основным принципом построения sc-языков является представление понятий, соответствующих основным классам объектов, описываемых sc-языком, и отношений между этими объектами ключевыми узлами такого sc-языка: каждому sc-языку однозначно сопоставляется конечное множество ключевых узлов (элементов) этого языка. Каждый ключевой узел задаёт ограничения на собственную семантическую окрестность в информационных конструкциях (текстах) этого языка. Множество ключевых элементов интегрированного sc-языка представления знаний является объединением множеств ключевых элементов остальных sc-языков выделенного семейства. Характеристиками sc-языка являются: мощность множества ключевых узлов sc-языка; семейство множеств собственных семантических окрестностей ключевых элементов языка; наличие функциональных зависимостей между собственными окрестностями множеств ключевых элементов sc-языка; соотношение алгоритмических сложностей поиска или вычисления элементов собственных семантических окрестностей на основании существующих зависимостей. Построена семантическая онтология существующих языков и моделей представления знаний, в которой указано место унифицированной модели представления знаний и её соотношение с существующими моделями представления знаний.

Введенные в sc-языки ключевые элементы поддерживают соответствующее семантически эквивалентное представление концептов OWL 2 QL и OWL 2 EL [W3C:OWL2, 2009], а также поддерживают представление множеств, мультимножеств, ситуативных множеств, отношений, включая ролевые отношения – строгие подмножества отношения принадлежности, простых, целых и рациональных чисел, логических формул. Остальные концепты могут быть определены с помощью логического sc-языка.

Модель унифицированного нелинейного представления знаний поддерживает представление знаний в условиях влияния различных НЕ-факторов: неполноты, неопределённости, нечёткости и пр. Далее приведены примеры представления нечётких и изменяющихся во времени знаний.

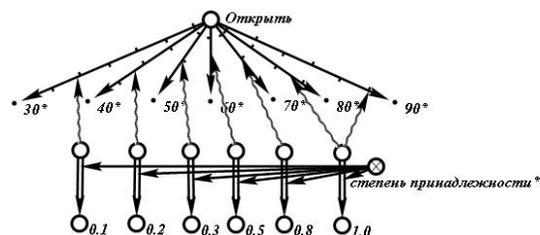


Рисунок 2– Пример представления нечёткого множества

Следующее нечёткое множество [Тэрano, 1993] описывает понятие открытия вентиля (рис. 2).

$$\text{Открыть} = 0,1/30^*+0,2/40^*+0,3/50^*+0,5/60^*+0,8/70^*+1,0/80^*+1,0/90^* \quad (12)$$

Степень нечёткой принадлежности указывается с помощью ключевого узла *степень принадлежности**, обозначающего бинарное отношение, связывающее sc-множество нечётких sc-дуг с числом, являющимся степенью принадлежности sc-элемента, в который входит нечёткая sc-дуга, sc-множеству, из которого она выходит. Следует отметить, что принадлежность нечётких sc-дуг классам с одинаковой степенью нечёткой принадлежности, в общем случае имеет нестационарную природу, что выражено соответствующими sc-дугами нестационарной принадлежности.

При представлении изменяющихся неточных и неопределённых знаний используются следующие языковые средства [Allen, 1991], [Allen, 1991] – временные отношения. В sc-языке временные отношения определяются на множестве sc-элементов, обозначающих нестационарные принадлежности и непринадлежности или их sc-множества, и выражают временные соотношения значений этих обозначений (знаков). Чтобы выразить временные отношения на самих обозначениях используется специальные ключевые узлы, связывающие значения одних обозначений, с обозначениями, временные отношения на которых необходимо выразить: такими узлами являются «полностью представленное множество», «sc-множество с полностью представленным экстенциональным замыканием» и другие.

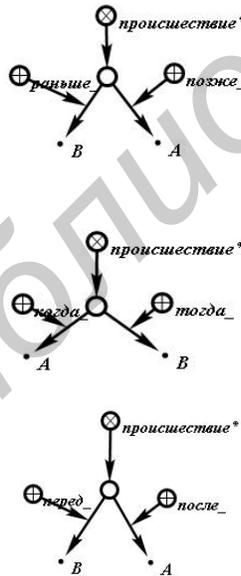


Рисунок 3– Примеры связей временных отношений.

Временное отношение, являющееся объединением всех неролевых временных отношений, обозначается ключевым узлом *происшествие**. Элементы связей этого отношения

отмечаются с помощью ролевых временных отношений (рис. 3). Временные отношения определяются в соответствии с отношением следования событий модели ситуативных множеств. Временное отношение, обозначаемое ключевым узлом *после*, выражает непосредственное следование всех событий одного множества за каждым событиями другого, к которым относится соответствующая принадлежность или непринадлежность или каждая из их sc-множества. Отношение *после* является иррефлексивным. Отношение, обозначенное ключевым узлом *перед*, является двойственным и обратным отношению *после*. Отношения *позже* и *раньше* соответственно выражают транзитивные замыкания отношений *после* и *перед*. Отношение *когда* выражает включение всех соответственных событий первого множества во второе множество событий или (в слабом случае) существование во втором множестве простых цепей событий, заданных отношением следования таких, что все события первого множества лежат на простых цепях, соединяющих начало какой-либо из этих цепей с её концом. Временное отношение *тогда* является двойственным отношению *когда*.



Рисунок 4– Пример представления временного отношения на нестационарных принадлежностях и заданных ими знаках множеств.

Рассмотрим представление знания (рис. 3) о том, что когда множество слева (обозначенное узлом с точкой) полностью представлено (в памяти системы представлены все обозначения его элементов и принадлежностей), то представлено и множество (обозначенное узлом с точкой), находящееся справа.

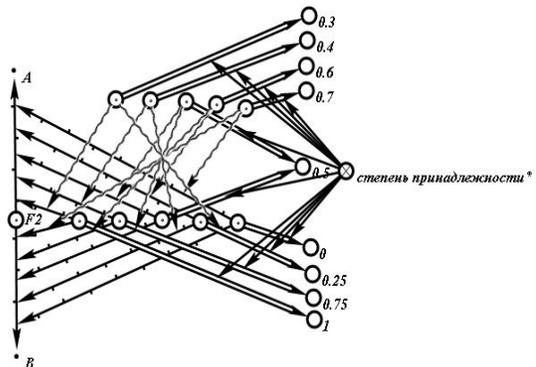


Рисунок 5– Пример представления нечёткого множества с нечётко заданными степенями принадлежности.

Для указания степеней нечёткой принадлежности элементов нечёткого множества второго порядка (2-типа) $F2$ (рис. 5), использованы пять дополнительных нечётких множеств первого порядка.

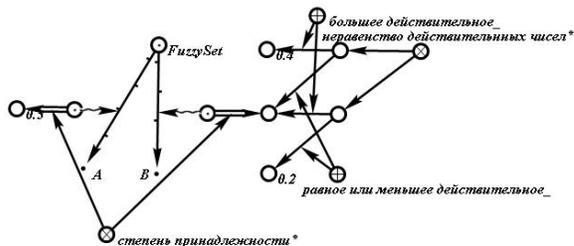


Рисунок 6– Пример представления нечёткого множества с неопределённой степенью принадлежности.

В случае, когда степень принадлежности точно указать нельзя, но можно оценить её значения сверху или снизу, для представления используются конструкции sc-языка арифметики. В нечётком множестве $FuzzySet$ (рис. 6), степень принадлежности которому элемента A известна, а степень принадлежности элемента B задана двумя числовыми рациональными оценками сверху и снизу.

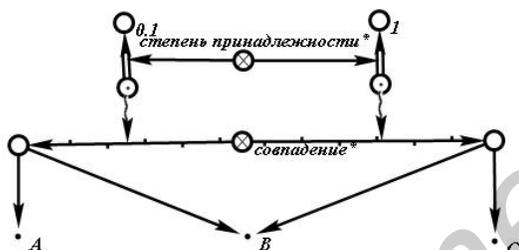


Рисунок 7– Пример представления нечёткого отношения.

Нечётко заданное отношение совпадения может использоваться при интеграции знаний в процессе выявления синонимов (рис. 7). Степень (нечёткой принадлежности связки отношения) совпадения sc-элемента A с sc-элементом B оценена в 0,1, а степень совпадения B с C – в единицу. Нечёткие sc-дуги принадлежности, имеющей степень равную единице, преобразуются в sc-дуги принадлежности, а те – которые имеют степень принадлежности равную нулю – в sc-дуги непринадлежности. Результат преобразования (рис. 8) есть текст, в котором обозначение принадлежности отмечено не как sc-дуга нечёткой принадлежности, а как sc-дуга принадлежности.

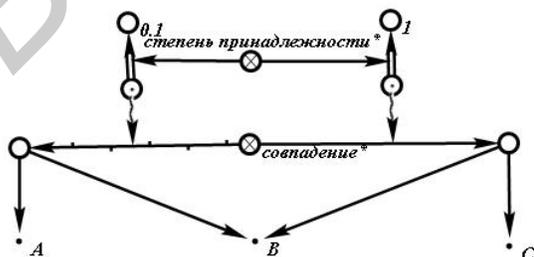


Рисунок 8– Пример перехода от нечёткого представления к чёткому.

В процессе рассуждений в интеллектуальной системе, нечёткая информация может уточняться. По мере того как происходят события исходная, оцененная в настоящий момент вероятность, сменяется условной по отношению к ней вероятностью уже в новый момент настоящего (рис. 9).

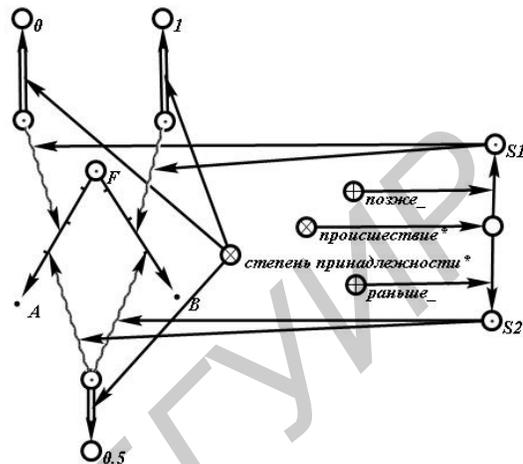


Рисунок 9– Пример представления изменяющегося во времени нечёткого множества.

В момент события, соответствующего узлу $S2$, степень принадлежности множеству F обоих sc-элементов A и B была оценена в 0,5, когда же наступило событие, соответствующее узлу $S1$, степень sc-элемента A стала равной нулю, а элемента B – единице.

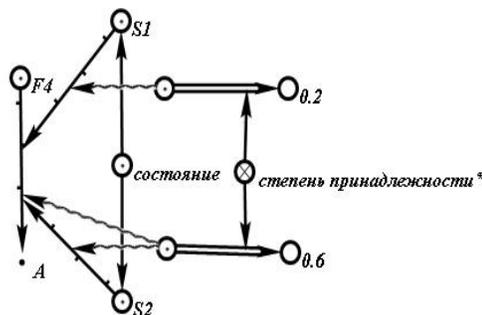


Рисунок 100– Пример представления множества с нечётко заданной ситуативностью.

Следует отметить, что сама информация о событиях, в которые рассматривается принадлежность или непринадлежность sc-элементов может быть нечёткой, как для нечёткого множества $F4$ (рис. 10). Элемент A принадлежит нестационарно множеству $F3$ (рис. 11), однако, то, что нечёткая оценка этой принадлежности соответствует событию, заданному узлом $S1$, равна 0,2, а – событию $S2$ – 0,6.

При интеграции знаний ситуативного (нестационарного) и гипотетического рода, каждое утверждение, которое содержится в исходных интегрируемых фрагментах, преобразуется по следующим схемам.

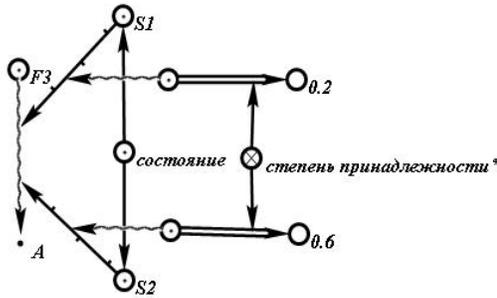


Рисунок 11– Пример представления нечёткого множества с нечётко заданной ситуативностью.

2. Интеграция знаний

Унифицированная модель баз знаний описывает базы знаний специального вида, в которых знания представлены с использованием унифицированной модели представления знаний [Ивашенко, 2009а]. Между базами знаний определены отношения содержательного и структурного включения одной базы знаний в другую. На множестве баз знаний заданы операции содержательной проекции и темпорализации [Ивашенко, 2009а]. Соответствие интеграции каждой паре баз знаний сопоставляет однозначное отображение множества всех знаков, формирующих эти базы знаний, на множество всех знаков результирующей базы знаний. Путём интеграции осуществляется переход от некоторой исходной базы знаний к требуемой оптимизированной базе знаний, которая имеет более высокое качество [Ивашенко, 2011а].

Решение задачи интеграции требуется в основных трёх случаях:

- интеграция разработанных и отлаженных баз знаний;
- интеграция отлаженных компонентов в базу знаний;
- добавление знаний при редактировании базы знаний.

В задаче интеграции двух баз знаний или их фрагментов в качестве исходных данных используются две БЗ (или фрагмента) и дополнительная метаинформация о свойствах знаков, принадлежащих этим базам знаний. Результатом решения этой задачи является нахождение интегрированной базы знаний, такой, что каждый знак из исходных интегрируемых баз знаний, имеет единственное представление в виде соответствующего знака в интегрированной базе [Ивашенко, 2009а]. Будем говорить, что осуществляется слияние двух знаков исходных БЗ тогда и только тогда, когда каждому из обоих соответствует в интегрированной базе знаний единственной знак. При интеграции в качестве дополнительной метаинформации может использоваться информация о внешних обозначениях (идентификаторах) понятий или информация, заданная базовой или ключевой рефлексивной семантикой знаков из этих БЗ. В

условиях неполноты информации можно выделить два типа стратегий слияния знаков интегрируемых баз знаний: безопасные (выполняемые однозначно и непротиворечиво) и небезопасные. Безопасная стратегия включается в любую стратегию слияния.

Задача интеграции баз знаний может быть разделена на две задачи – интеграция содержания баз знаний и интеграция структуры баз знаний (обеспечение целостности базы знаний).

Без потери общности для решения задачи интеграции содержания баз знаний рассмотрим задачу интеграции двух фрагментов баз знаний, каждый из которых представляет связную информационную конструкцию sc-языка.

При интеграции двух баз знаний между ними выявляется некоторое множество пар потенциально синонимичных элементов. Методика точного решения этой задачи включает как решение этой задачи вручную, путём получения соответствующего ответа от разработчика, так и автоматизированное решение на основе известной базовой теоретико-множественной семантики элементов и ключевых узлов sc-языка. Исходя только из базовой теоретико-множественной интерпретации, уже в некоторых фрагментах базы знаний можно легко установить факт наличия или отсутствия синонимии. Зная множество таких пар и множество всех возможных пар всех элементов в заданном фрагменте базы знаний, легко вычислить множество и число пар потенциально синонимичных элементов такой базы знаний.

Имея информацию о потенциально синонимичных элементах, можно оценить множество и количество всевозможных структурно различимых вариантов слияния потенциально синонимичных элементов в элементы результирующей базы знаний, являющееся подмножеством результатов соответствия интеграции. Это множество обозначим величиной $I(G)$, где G – симметричный ориентированный граф, множество рёбер $E(G)$ которого является множеством всех пар потенциально синонимичных элементов исходных баз знаний, а множество вершин $V(G)$ – множеством всех элементов исходных баз знаний

$$E(G) \subseteq (V(G))^2. \quad (13)$$

Для того, чтобы определить множество $I(G)$, используем специальную операцию

$$A \overset{\text{def}}{\cup} B = \bigcup_{(P,Q) \in A \times B} \{P \cup Q\}, \quad (14)$$

которая является ассоциативной и коммутативной.

Чтобы задать множество $I(G)$, введём семейство всевозможных множеств рёбер ориентированных графов $C(G)$ на множестве вершин графа G , каждая компонента связности которых является полным подграфом [Ивашенко, 2009b], тогда

$$I(G) = \left(\bigcup_{e_{ij} \in E(G)} \{\emptyset, \{e_{ij}\}\} \right) \cap C(G). \quad (15)$$

Число Q элементов множества $I(G)$ можно рассматривать в виде критерия качества: чем меньше это число, тем выше качество и наоборот. Однако, уже для небольших фрагментов баз знаний, это число может оказаться большим и трудным для расчёта, поэтому на практике более целесообразно использовать или логарифм этого числа или логарифм его оценок, или связанные с ними величины (число пар потенциальных синонимов).

$$Q = |I(G)|. \quad (16)$$

В работе [Ивашенко, 2009b] для числа Q установлены соотношения для верхней и нижней границы.

Интеграция онтологий и баз знаний осуществляется через слияние знаков. Рассмотрим m объектов области значений семантической интерпретации некоторого знака. Тогда, на множестве этих объектов, число неоднозначных и однозначных постоянных семантических интерпретаций (семантик) знака равно $2^{2^m} - 1$.

Чтобы осуществить слияние знаков (воплощений знака) необходимо, чтобы пересечение областей значения их семантик было непустым. В частности – теоретико-множественных семантик.

Если пересечение областей значения семантик знаков является пустым, то слияние таких знаков невозможно и такие знаки различны.

Тогда как для выявления различных знаков достаточно одного контр-примера, для выявления подлежащих слиянию знаков необходимо не только установить, что пересечение областей значения их семантик не является пустым, но и доказать, что эти знаки совпадают. Для этой цели используются различные утверждения о совпадении в виде аксиом и теорем.

Опишем свойства этого отношения совпадения. Если два знака совпадают, то не существует множества, которому одновременно принадлежит один из них и не принадлежит другой. Если два знака множеств совпадают, то множества равны. Выявить неравенство множеств просто: достаточно найти элемент, который принадлежит одному множеству и не принадлежит другому. Процедура выявления равенства множеств зависит от их свойств. Если множества имеют конечное, не очень большое число принадлежностей или непринадлежностей элементов, что характерно для понятий, то перечислив все принадлежности или все непринадлежности, число которых совпадает соответственно с его мощностью или количеством непринадлежностей, и убедившись, что среди них нет непринадлежностей элементов, принадлежащих другому множеству, можно заключить, что множества равны. Если же множества бесконечны или число принадлежностей, равно как и непринадлежностей у них, очень большое, то

равенство множеств можно доказать только через утверждения об их свойствах.

Прежде, чем проводить слияние знаков двух баз знаний, следует провести отображение баз знаний (по аналогии с отображением онтологий).

Для отображения в некоторых работах [Maltese et al., 2010] используются следующие отношения между онтологическими понятиями $\{\equiv, \supseteq, \sqsubseteq, \perp, \sqcap\}$, соответственно – совпадения, обобщения, частности, строгого исключения и строгого пересечения. Например, эти отношения в проекции на теоретико-множественные соответствуют отношениям равенства множеств ($=$), надмножества (\supseteq), подмножества (\sqsubseteq), пустого пересечения множеств ($S1 \cap S2 = \emptyset$) и непустого пересечения множеств ($S1 \cap S2 \neq \emptyset$). Однако, как можно показать, эти отношения трудно или невозможно установить, когда множества имеют большую или бесконечную мощность, или если исходить из того, что существуют равные, но несовпадающие множества.



Рисунок 12– Пример, приводящий к противоречивому отображению

Для приведённого примера (рис. 12), при условии известности, что понятия ромба и квадрата не пересекаются, будет получено противоречивое «сильное» отображение (многоугольник (S1), многоугольник \wedge четырёхугольник (S2), многоугольник \wedge четырёхугольник \wedge ромб (S3), многоугольник \wedge четырёхугольник \wedge квадрат (S4), замкнутая ломаная (S5), замкнутая ломаная \wedge четырёхугольник (S6), замкнутая ломаная \wedge четырёхугольник \wedge прямоугольник (S7), замкнутая ломаная \wedge четырёхугольник \wedge параллелограмм (S8), замкнутая ломаная \wedge четырёхугольник \wedge параллелограмм \wedge ромб (S9)).

Таблица 1 – Противоречивое отображение

	S1	S2	S3	S4
S5	\equiv	\supseteq	\supseteq	\supseteq
S6	\sqsubseteq	\equiv	\supseteq	\supseteq
S7	\sqsubseteq	\sqsubseteq	\equiv	\equiv
S8	\sqsubseteq	\sqsubseteq	\equiv	\supseteq
S9	\sqsubseteq	\sqsubseteq	\equiv	\perp

Кроме этого предложенные [Maltese et al., 2010] методы отображения применимы к так называемым «легко-взвешенным» онтологиям [Giunchiglia et al., 2006] и не применимы к другим, аналогичные трудности встречаются и в других подходах [Aumueller et al., 2005], [Jean-Mary et al., 2007], [Nagy et al., 2010].

Исходя из этого, автором предложен следующий набор отношений {не уточнены*, связность*, различие*, исключение* (исключаемое_, исключающее_), исключение пересекающихся* (исключаемое пересекающееся_, исключающее пересекающееся_), симметричное исключение*, строгое пересечение*, совпадение*}; отношение потенциальной синонимии можно выразить, как $возможная\ синонимия^* = (не\ уточнены^* \cup связность^*) / (исключение^* \cup исключение\ пересекающихся^* \cup различие^*)$.

Следующее правило (рис. 13) позволяет в рамках унифицированной модели представления знаний единообразно, непротиворечиво и строго, что было одной из трудностей в [Maltese et al., 2010], задать правила отображения и слияния знаков баз знаний при использовании методов лексико-терминологического анализа (анализа идентификаторов).

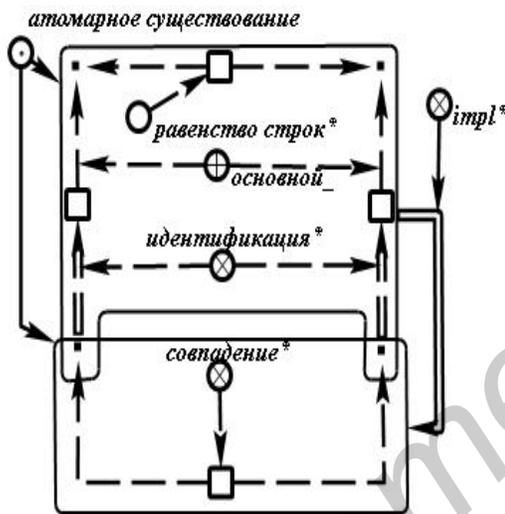


Рисунок 13 – Пример, приводящий к противоречивому отображению

Стратегии слияния при использовании методов лексико-терминологического анализа подробно рассмотрены в [Ивашенко, 2011b].

Следующие правила (рис. 14-19) описывают установление отображающих отношений на множестве знаков интегрируемых фрагментов баз знаний, после проведённых слияний в результате лексико-терминологического анализа.

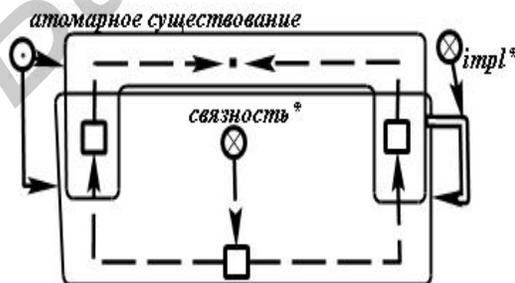


Рисунок 14 – Свойство отношения связности («два множества связны, если имеют общий элемент»)

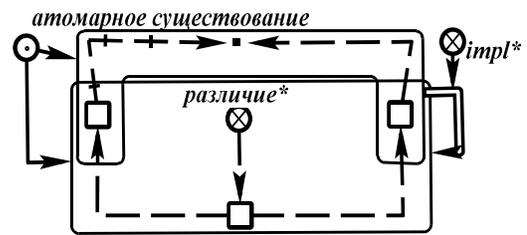


Рисунок 15 – Свойство отношения различия («два элемента различны, если один принадлежит множеству, а другой – нет»)

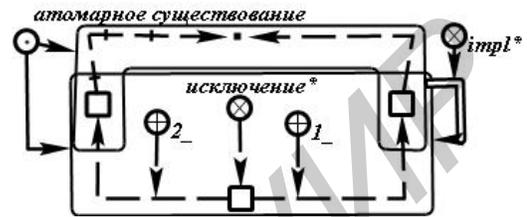


Рисунок 16 – Свойство отношения исключения (исключаемое_, исключающее_)

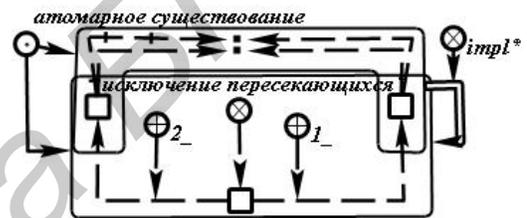


Рисунок 17 – Свойство отношения исключения пересекающихся

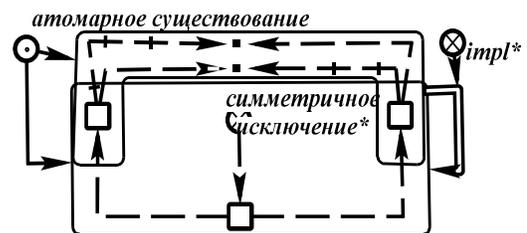


Рисунок 18 – Свойство отношения симметричного исключения

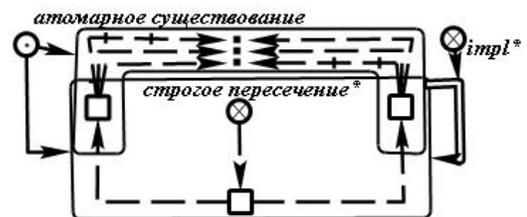


Рисунок 19 – Свойство отношения строгого пересечения

Приведём некоторые правила, описывающие взаимные свойства отношений (рис. 20-24).

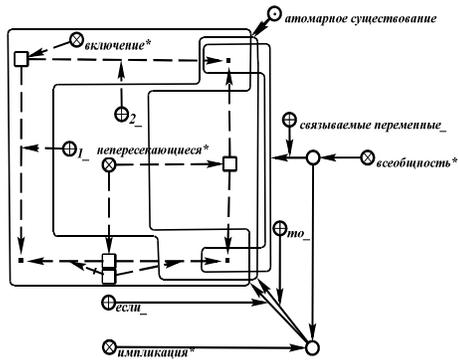


Рисунок 20 – Свойство отношения непересечения

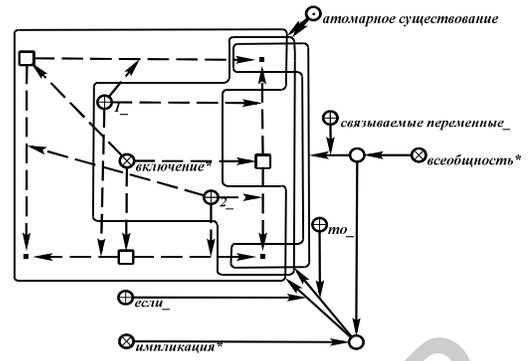


Рисунок 23 – Свойство транзитивности отношения включения

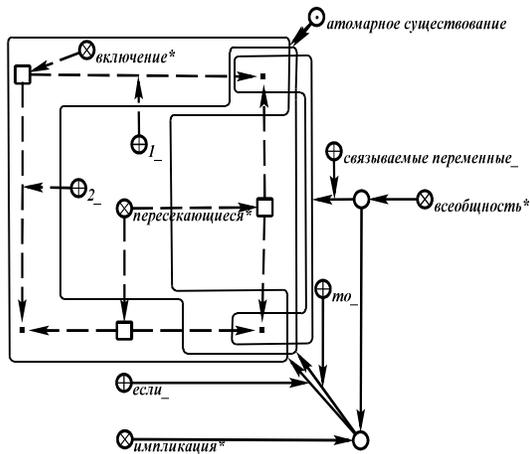


Рисунок 21 – Свойство отношения пересечения

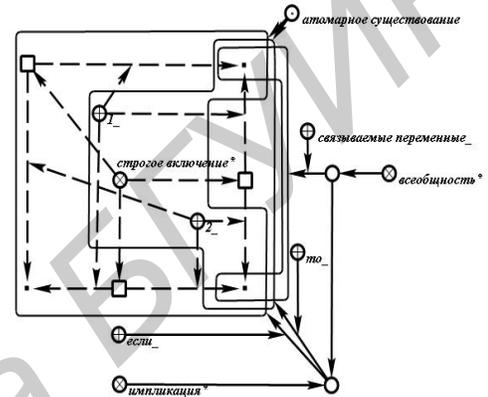


Рисунок 24 – Свойство транзитивности отношения строгого включения

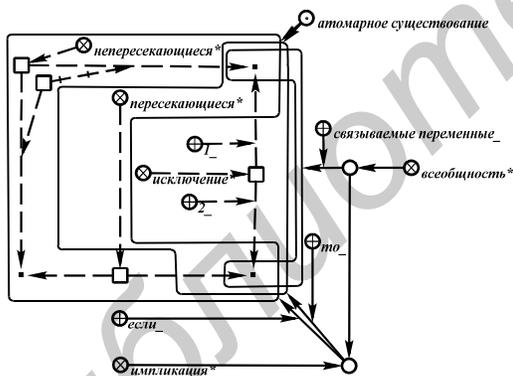


Рисунок 22 – Свойства отношения исключения

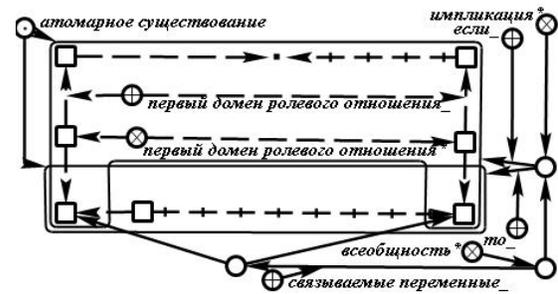
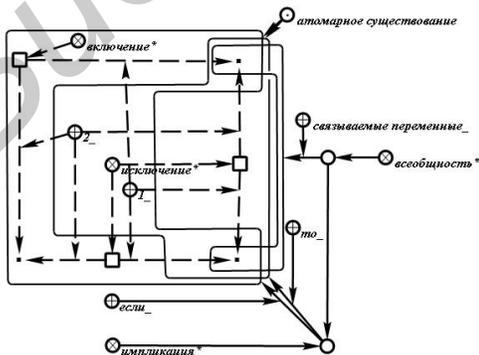


Рисунок 25 – Свойство ролевых отношений с различными первыми доменами

Для учёта теоретико-множественных отношений объединения и пересечения и (относительного) дополнения (как это предлагается в S-match) в данной задаче достаточно, например, добавить такие отношения как относительное дополнение (разность) и семейство непересекающихся множеств (двух и более) и задать их соответствующие свойства по отношению к ранее введённым отношениям.

При интеграции реляционных моделей можно придерживаться следующего порядка действий:

- интеграция ролевых отношений и их связок (рис. 25, 26),
- интеграция неролевых отношений (рис. 27, 28).

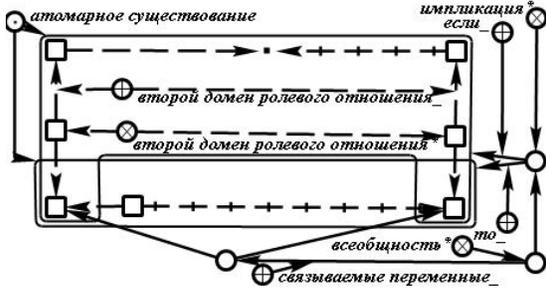


Рисунок 26 – Свойство ролевых отношений с различными вторыми доменами

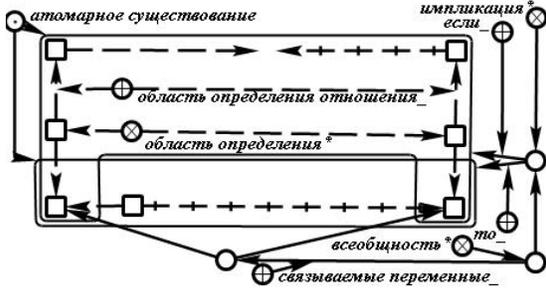


Рисунок 27 – Свойство отношений с различными областями определения

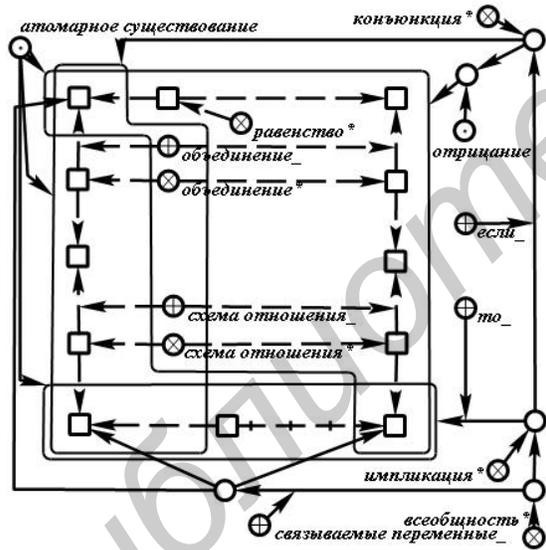


Рисунок 28 – Свойство отношения с различными объединениями схем

Важным признаком совпадения элементов является их принадлежность одной связке под однозначным ролевым отношением или наличие принадлежности им общего элемента под обратно однозначным ролевым отношением.

Все эти отношения можно однозначно и непротиворечиво установить в результате соответствующего базовой теоретико-множественной семантике структурного анализа, включая экстенциональный, за

полиномиальное время. Более глубокий логико-семантический анализ может потребовать больших временных затрат.

Для анализа на логико-семантическом уровне важно выделять утверждения, которые приводят к совпадению знаков интегрируемых фрагментов, в результате чего можно произвести их слияние. Такowymi высказываниями являются высказывания о свойствах логических высказываний о единственности, высказывания о свойствах множеств без кратных вхождений элементов (канторовских множеств), высказывания об отношениях без кратных связок. Ниже приведён пример (рис. 29), описывающий свойство логических утверждений о единственности и сводящий задачу анализа таких утверждений к задаче выявления связок отношений совпадения. Таким образом, при наличии средств логического вывода для решения задачи интеграции логико-семантическими методами достаточно описать свойства соответствующих ключевых узлов, выражающих количественные ограничения (например, единственность) или описывающих отсутствие кратных связок, через понятия совпадения sc-элементов (знаков текстов sc-языка).

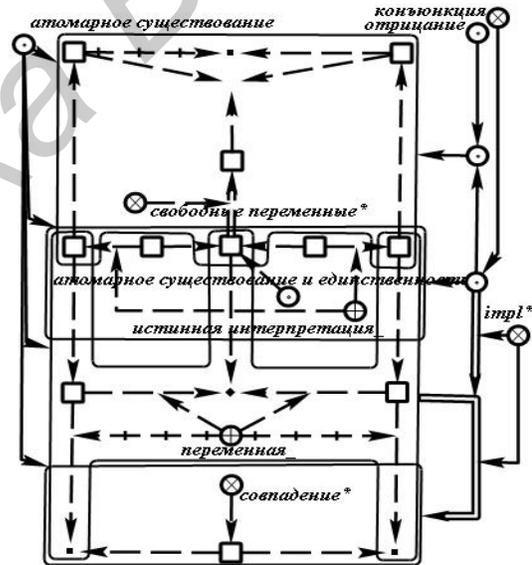


Рисунок 29 – Высказывание о слиянии значений переменных в утверждениях единственности

Для интеграции фрагмента базы знаний требуются следующие sc-операции, решающие следующие задачи:

- поиск и выборка элементов заданного интегрируемого фрагмента;
- проверка и выборка элементов, имеющих строковое содержимое и хранящих внешние обозначения (идентификаторы);
- проверка и выборка элементов, имеющих основной идентификатор;
- проверка и выборка элементов, имеющих неосновной идентификатор;

- добавление в словарь элемента с идентификатором;
- поиск и выборка элементов из словаря по идентификатору;
- формирование для заданных идентификаторов факта одинаковых идентификаторов;
- слияние совпадающих элементов;
- выявление и формирование фактов потенциально синонимичных элементов;
- выявление и формирование отображения знаков онтологий;
- формирование исходного варианта слияния элементов заданных разделов;
- применение к заданным элементам утверждений о совпадении элементов, в частности элементов, имеющих одинаковый основной идентификатор, утверждений о различии (несовпадении) и о возможном совпадении элементов;
- выбор и формирование варианта слияния заданных элементов для установленного факта возможного совпадения и заданного варианта слияния;
- выбор варианта слияния элементов базы знаний, исключение невыбранных вариантов слияния.

Обобщённый алгоритм интеграции фрагментов двух баз знаний приведён в [Ивашенко, 2011b]. Ниже (рис. 30) приведён алгоритм выявления пар потенциальных синонимов, использующий вышеописанные правила экстенционального метода.

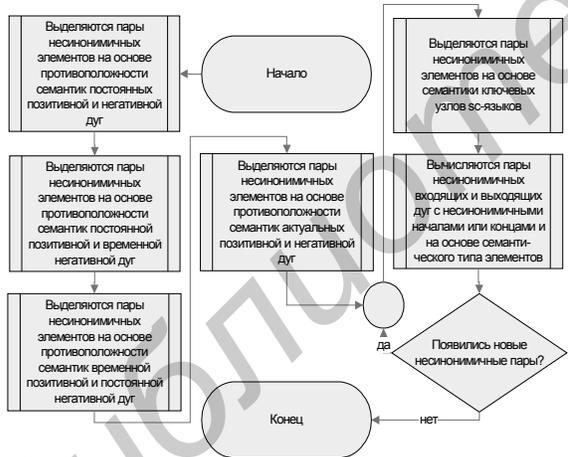


Рисунок 30 – Алгоритм вычисления множества пар различающихся (несинонимичных) элементов

Интеграция структуры баз знаний может быть проведена на исходном уровне или на метуровне. В первом случае, если в базе знаний после содержательной интеграции появились разделы, слияние которых не произошло, то формируется новый раздел базы знаний, который декомпозируется на основные (недекомпозируемые) разделы интегрируемых баз знаний. Этот вариант допустим, если декомпозиция разделов сохраняет древовидную структуру. Вторым вариантом предполагает формирование мета-разделов, которые описывают основные разделы исходных баз

знаний и их структуру и на которые декомпозируется основной раздел интегрированной базы знаний. Вторым вариантом интеграции структуры баз знаний может иметь смысл, когда содержание исходных баз знаний является слабо связным.

При интеграции тезаурусов необходимо построить связную иерархию определяемых понятий. Для каждого понятия необходимо указать является оно определяемым или нет. Для каждого понятия необходимо указать множества понятий, на основе которых оно может быть однозначно определено, а на основе каких – нет. Для построения такой иерархии применим алгоритм, подобный обобщённому алгоритму интеграции. При поиске выполнимых интерпретаций схем обход возможных предикативных выражений осуществляется по всевозможным предварённым конъюнктивным нормальным формам, атомарные подформулы которых соответствуют алфавитным меткам, связкам инцидентности и равенству значения переменной константе и задают всевозможные размеченные привязанные с помощью констант подграфы в графе онтологии. Для сокращения перебираемых вариантов используются эвристики.

Интеграция тезаурусов (баз знаний, содержащих определения) может быть осуществлена на основе отсеивания наборов понятий, которые не могут относиться как определяющие и определяемые, в результате анализа по следующим схемам (фильтрация неоднозначно определяющих множеств понятий для каждого понятия).

Схема n-значности ($|v| = n$) свойства α задана следующим выражением.

$$\lambda(\alpha, x, v, z) = (\alpha(\langle x, v, z \rangle) \wedge \forall w (\alpha(\langle x, w, z \rangle) \sim (w \in v))) \quad (17)$$

Схема необходимости.

$$\neg \exists U \exists V \exists Y (\neg \beta(\langle X, U, Y \rangle) \wedge \gamma(\langle Y, V, Z \rangle)) \quad (18)$$

Схема достаточности.

$$\neg \exists U \exists V \exists Y (\beta(\langle X, U, Y \rangle) \wedge \neg \gamma(\langle Y, V, Z \rangle)) \quad (19)$$

Схема прямой однозначности.

$$\exists U \exists V \exists Y (\lambda(\langle \beta, X, \{U\}, Y \rangle) \wedge \neg \lambda(\langle \gamma, Y, \{V\}, Z \rangle)) \quad (20)$$

Схема обратной однозначности.

$$\exists U \exists V \exists Y (\lambda(\langle \gamma, Y, \{V\}, Z \rangle) \wedge \neg \lambda(\langle \beta, X, \{U\}, Y \rangle)) \quad (21)$$

Здесь U, V, X, Y, Z являются множествами элементов: X – определяющий набор понятий; Z – определяемый набор понятий; U, V, Y – элементы окрестности X и Z ; α, β и γ – предикативные выражения определяющих и определяемых свойств.

Рассмотрим пример интеграции двух фрагментов баз знаний (см. рис. 31).

Для решения приведённой в примере задачи интеграции требуется следующая дополнительная информация:

- информация о равенстве основных идентификаторов;
- информация о равенстве неосновных идентификаторов;
- информация о том, какие множества являются множествами без кратных вхождений элементов (канторовские множества);
- информация об отсутствии кратных связей в геометрических отношениях;
- утверждение об однозначном задании биссектрисы заданного треугольника, проходящей через заданную вершину;
- утверждение об однозначном задании треугольника заданными тремя вершинами
- утверждение об однозначном задании вписанной окружности заданным треугольником;
- утверждение о совпадении центра вписанной в треугольник окружности и точки пересечения биссектрис треугольника.

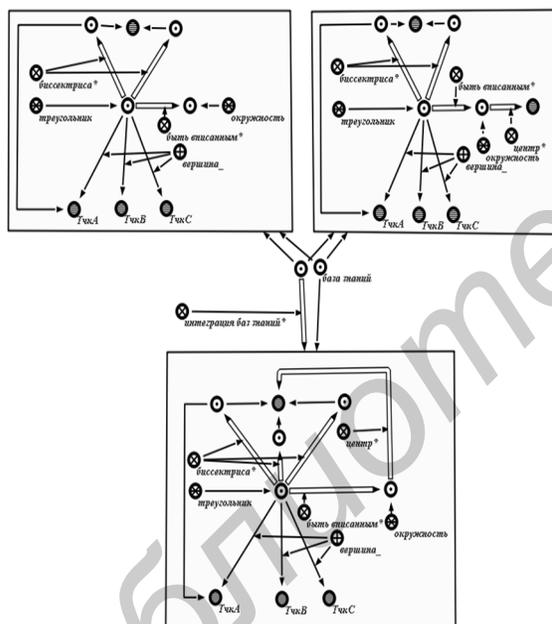


Рисунок 31 – Пример интеграции двух фрагментов баз знаний

ЗАКЛЮЧЕНИЕ

Рассмотренные модели и средства в рамках семантической технологии проектирования баз знаний [OSTIS, 2010] обеспечивают разработку баз знаний путём добавления и интеграции [Гулякина и др., 2004] в разрабатываемую базу знаний многократно используемых компонентов баз знаний, что позволяет сократить сроки её разработки.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [Гаврилова и др., 2000] Гаврилова, Т.А. Базы знаний интеллектуальных систем / Т. А. Гаврилова, В.Ф. Хорошевский. – СПб.: Питер, 2000.
- [Голенков и др., 2001] Голенков, В.В. Представление и обработка знаний в графодинамических ассоциативных машинах / В. В. Голенков[и др.] – Мн. : БГУИР, 2001.
- [Гулякина и др., 2004] Гулякина Н.А. Ивашенко В.П. Интеграция знаний в информационных системах. / Н.А. Гулякина, В.П. Ивашенко // Доклады БГУИР. – 2004. – №6. – С. 113-119.
- [Ивашенко, 2003] Представление нейронных сетей и систем продуцций в однородных семантических сетях. / В.П. Ивашенко // Известия Белорусской инженерной академии. – 2003. – №1(15)/1. – с.184-188.
- [Ивашенко, 2004] Применение однородных семантических сетей для представления знаний о нестационарных предметных областях. / В.П. Ивашенко // Известия Белорусской инженерной академии. – 2004. – №1(17)/3. – с.77-80.
- [Ивашенко, 2009a] Ивашенко В.П. Семантические модели баз знаний / В.П. Ивашенко Информационные системы и технологии (IST'2009): материалы V Междунар. конф.-форума в 2-х ч. Ч.2 – Минск: А.Н.Вараксин, 2009. - с.125-128.
- [Ивашенко, 2009b] Ивашенко В.П. Алгоритмы верификации и интеграции баз знаний. Вестник Брестского государственного технического университета. БрГТУ, 2009, №5.
- [Ивашенко, 2011a] Ивашенко В.П. Семантическая технология компонентного проектирования баз знаний. Материалы Международной науч.-техн. Конференции OSTIS,2011:Минск, Республика Беларусь, БГУИР 10-12 февраля 2011.
- [Ивашенко, 2011b] Ивашенко В.П. Алгоритмы операций отладки и интеграции баз знаний. Дистанционное обучение – образовательная среда XXI века: материалы VII Международной науч.-метод. конференции (Минск, 1-2 декабря 2011г.). – Минск: БГУИР, 2011, сс.227-229.
- [Ивашенко, 2012] Ивашенко В.П. Семантическая модели интеграции и отладки баз знаний. Материалы Международной науч.-техн. Конференции OSTIS,2012:Минск, Республика Беларусь, БГУИР 16-18 февраля 2012.
- [Колб, 2012] Колб, Д. Г. Web-ориентированная реализация семантических моделей интеллектуальных систем / Д. Г. Колб // Материалы международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» – Минск, 2012
- [Кофман, 1982] Кофман А. Введение в теорию нечетких множеств. М.: Радио и связь, 1982. — 432 с.
- [Кудрявцев, 2008] Кудрявцев Д. В. Практические методы отображения и интеграции онтологий. Семинар Знания и онтологии *Elsewhere*. КИИ-2008, Дубна, 2008.
- [Нариньяни, 2000] Нариньяни А. С. НЕ-факторы: неточность и недоопределенность — различие и взаимосвязь // Изв. РАН. Сер. Теория и системы управления. 2000. № 5. С. 44-56.
- [Тэрано, 1993] Прикладные нечеткие системы: Пер. с япон. / К. Асаи, Д. Ватада, С. Иваи и др.; Под ред. Т. Тэрано, К. Асаи, М. Сугэно. – М.: Мир, 1993.
- [Хорошевский, 2008] Хорошевский, В.Ф. Пространства знаний в сети Интернет и Semantic Web (Часть 1) / В. Ф. Хорошевский // Искусственный интеллект и принятие решений. - 2008. - № 1. - С.80-97.
- [Allen, 1991] Allen, J.F.: Time and time again: the many ways to represent time. International Journal of Intelligent Systems 6 (1991) 341-355.
- [Allen, 1994] Allen, J. F. Ferguson, G. Actions and Events in Interval Temporal Logic. J. Log. Comput., 1994, 4, 531-579.
- [Aumüller et al., 2005] D. Aumüller, H. Do, S. Massmann, E. Rahm. 2005. Schema and ontology matching with COMA++. Proc. of the 2005 International Conference on Management of Data, pp. 906-908.
- [Bouquet, 2003] P. Bouquet, L.Serafini and S. Zanobini, M. Benerecetti. An algorithm for semantic coordination. Proc. of Semantic Integration Workshop, collocated with the 2nd Int.l Semantic Web Conference (ISWC2003), 20-23 October 2003, Sanibel Island, Florida, US.
- [DAML, 2006] DAML.org: The DARPA Agent Markup Language Homepage, Электронный ресурс. Режим доступа: <http://www.daml.org/>.
- [Doan and Halevy, 2005] A. H. Doan, A. Y. Halevy. Semantic integration research in the database community: A brief survey. AI magazine, 26(1), 2005

- [**CLIPS, 1991**] CLIPS Version 5.1 User's Guide, NASA Lyndon B. Johnson Space Center, Software Technology Branch, Houston, TX, 1991.
- [**CoFI:CASL-Summary, 2004**] CoFI Language Design Group. Casl summary. In Casl Reference Manual, LNCS 2960 (IFIP Series), part I. Springer, 2004.
- [**CycL, 2002**] The Syntax of CycL. Электронный ресурс. Режим доступа: <http://www.cyc.com/cycdoc/ref/cycl-syntax.html>.
- [**Ganter, 1999**] B. Ganter, R. Wille: Formal Concept Analysis – Mathematical Foundations. Springer, Heidelberg 1999.
- [**Gangemi et al., 1996**] A. Gangemi, G. Steve and F. Giacomelli, 1996. "ONIONS: An Ontological Methodology for Taxonomic Knowledge Integration", ECAI-96 Workshop on Ontological Engineering, Budapest, August 13th.
- [**Genesereth and Fikes, 1992**] Michael R. Genesereth and Richard E. Fikes. Knowledge interchange format version 3.0 reference manual. Logic Group First Draft January 1992 Report Logic-92-1 Current Version June 1992.
- [**Giunchiglia et al., 2006**] F. Giunchiglia, M. Marchese and I. Zaihayeu (2006). Encoding classifications into lightweight ontologies. University of Trento Technical Report # DIP-06-016, March 2006.
- [**Jean-Mary et al., 2007**] Jean-Mary Y., Kabuka, M. ASMOV: Ontology Alignment with Semantic Validation. Joint SWDB-ODDBIS Workshop, September 2007, Vienna, Austria, 15-20.
- [**IEEE, 2006**] IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, Automatic Fuzzy Ontology Generation for Semantic Web, VOL. 18, NO. 6, JUNE 2006.
- [**IDEF5, 1994**] Information Integration for Concurrent Engineering (ICE). IDEF5 Method Report. - Knowledge Based Systems, Inc. , 1408 University Drive East College Station, Texas, USA. September 21. 1994.
- [**ISO13250**] ISO/IEC 13250:2003 Topic maps – representation and interchange of knowledge, with an emphasis on the findability of information.
- [**ISO24707**] ISO/IEC 24707:2007 – Information technology – Common Logic (CL): a framework for a family of logic-based languages.
- [**Maldavan, 2001**] J. Madhavan, P. A. Bernstein, E. Rahm. Generic Schema Matching with Cupid. In. Proc. of the 27th Conference on Very Large Databases, 2001.
- [**Maltese et al., 2010**] V. Maltese, F. Giunchiglia, A. Autayeu: Save up to 99% of your time in mapping validation In Proceedings of ODBASE, 2010.
- [**Martin Ph., 2002**] Martin Ph. (2002). Knowledge representation in CGLF, CGIF, KIF, Frame-CG and Formalized-English. Proceedings of ICCS 2002, 10th International Conference on Conceptual Structures (Springer, LNAI 2393, pp. 77-91), Borovets, Bulgaria, July 15-19, 2002.
- [**McGuinness et al., 2000**] D. L. McGuinness, R. Fikes, J. Rice, and S. Wilder. The chimaera ontology environment. In Proceedings of AAAI, pages 1123-1124, 2000.
- [**Mitra, 2001**] Prasenjit Mitra, Gio Wiederhold and Stefan Decker: A Scalable Framework for Interoperation of Information Sources. 1st International Semantic Web Working Symposium (SWWS '01), Stanford University, Stanford, CA, July 29-Aug 1, 2001, Jul.
- [**Michael Kifer et al., 1995**] Michael Kifer, Georg Lausen, James Wu: Logical Foundations of Object-Oriented and Frame-Based Languages. Journal of ACM 42(4): 741-843 May (1995).
- [**Mustafa Jarrar et al., 2008**] Mustafa Jarrar and Robert Meersman: "Ontology Engineering -The DOGMA Approach". Book Chapter (Chapter 3). In Advances in Web Semantics I. Volume LNCS 4891, Springer, 2008.
- [**Nagy et al., 2010**] Miklos Nagy and Maria Vargas-Vera. Towards an Automatic Semantic Data Integration: Multi-Agent Framework Approach. Sematic Web. Gang Wu (ed), Chapter 7, pp. 107-134; In-Tech Education and Publishing KG; 2010, ISBN 978-953-7619-54-1.
- [**Novak et al., 2008**] Joseph D. Novak & Alberto J. Canas, The Theory Underlying Concept Maps and How to Construct and Use Them. Florida Institute for Human and Machine Cognition. Pensacola Fl, 32502. Technical Report IHMC CmapTools 2006-01 Rev 2008-01.
- [**OSTIS, 2010**] Открытая семантическая технология проектирования интеллектуальных систем [Электронный ресурс]. – 2010. - Режим доступа: <http://ostis.net>. – Дата доступа: 2.11.2010
- [**Smith et al., 2007**] Smith, B.; Ashburner, M.; Rosse, C.; Bard, J.; Bug, W.; Ceusters, W.; Goldberg, L. J.; Eilbeck, K. et al. (2007). "The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration". Nature Biotechnology 25 (11): 1251-1255. doi:10.1038/nbt1346.
- [**Sowa et al., 2008**] Sowa, J. Conceptual Graphs/ John F. Sowa, F. van Harmelen, V. Lifschitz, B. Porter// eds., Handbook of Knowledge Representation, Elsevier, 2008, pp. 213-237.
- [**Stumme, 2001**] Gerd Stumme and Alexander Maedche. FCA-merge: bottom-up merging of ontologies. In In Proceedings of 17th IJCAI, pages 225{230, Seattle (WA), USA, 2001.
- [**W3C:DAML+OIL, 2001**] DAML+OIL (March 2001) Reference Description: W3C Note 18 December 2001. Dan Connolly, Frank van Harmelen, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, Lynn Andrea Stein and Lucent Technologies, Inc, Электронный ресурс. Режим доступа: <http://www.w3.org/TR/daml+oil-reference>.
- [**W3C:RDFS, 2004**] RDF Vocabulary Description Language 1.0: RDF Schema: W3C Recommendation 10 February 2004 Электронный ресурс. / eds. Dan Brickley, R.V. Guha. Режим доступа: <http://www.w3.org/TR/rdf-schema/>, свободный.
- [**W3C:RIF, 2010**] RIF Overview: W3C Working Group Note 22 June 2010. Электронный ресурс. / eds. Michael Kifer, Harold Boley. Режим доступа: <http://www.w3.org/TR/2010/NOTE-rif-overview-20100622/>.
- [**W3C:OWL, 2004**] OWL Web Ontology Language. Overview: W3C Recommendation 10 February 2004 Электронный ресурс. / eds. Deborah L. McGuinness, Frank van Harmelen. Режим доступа: <http://www.w3.org/TR/owl-features/>.
- [**W3C:OWL2, 2009**] OWL 2 Web Ontology Language Document Overview: W3C Recommendation 27 October 2009. Электронный ресурс. / eds. W3C OWL Working Group. Режим доступа: <http://www.w3.org/TR/owl2-overview/>.
- [**W3C:SPARQL, 2008**] SPARQL Query Language for RDF: W3C Recommendation 15 January 2008. Электронный ресурс. / eds. Eric Prud'hommeaux, Andy Seaborne. Режим доступа: <http://www.w3.org/TR/rdf-sparql-query/>.
- [**XTM, 2001**] XML Topic Maps (XTM) 1.0: TopicMaps.Org Specification. Members of the TopicMaps.Org Authoring Group. Электронный ресурс. / eds. Steve Pepper, Graham Moore. Режим доступа: <http://www.topicmaps.org/xtm/>.
- [**Van Renssen, 2005**] Van Renssen, Andries (2005). Gellish: A Generic Extensible Ontological Language. Delft University Press. ISBN90-407-2597-4. <http://repository.tudelft.nl/view/ir/uuid%3Ade26132b-6f03-41b9-b882-c74b7e34a07d/>.

UNIFIED KNOWLEDGE REPRESENTATION AND KNOWLEDGE INTEGRATION

Ivashenko V.P.

*Belarusian State University of Informatics and
Radioelectronics, Minsk, Republic of Belarus*

ivashenko@bsuir.by

The work concerns models for unified knowledge representation and knowledge integration in the component knowledge base design technology (OSTIS). Several schemas and rules for algorithms of knowledge and ontology integration are considered. These rules and schemas provide ontology integration in conditions of dynamically updating knowledge and knowledge base incompleteness.