



OSTIS-2014

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

О ФОРМАЛИЗАЦИИ СЕМАНТИКИ ОБЛАСТЕЙ ЗНАНИЙ В ИНФОРМАЦИОННЫХ И ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ НА ОСНОВЕ ОНТОЛОГИЙ

Загорулько Ю.А. *, Загорулько Г.Б. *

** Институт систем информатики им. А.П. Ершова Сибирского отделения
Российской академии наук, г. Новосибирск, Россия*

zagor@iis.nsk.su

gal@iis.nsk.su

В докладе рассматриваются подходы и средства формализации семантики областей знаний на основе онтологий, обсуждается роль онтологий в интеллектуальных и информационных системах, обсуждаются примеры практического использования онтологий в системах такого класса. В заключении делается вывод, что онтологий могут успешно использоваться как при проектировании и разработке информационной и интеллектуальной системы, так и в качестве полноправного компонента во время ее функционирования.

Ключевые слова: формализация семантики; онтология; интеллектуальная система; база знаний.

Введение

Для того чтобы компьютер мог стать интеллектуальным помощником человека в его повседневной деятельности, необходимо формализовать и семантику тех аспектов этой деятельности, которые мы хотим автоматизировать, и семантику предметной области (ПО), в рамках которой эта деятельность осуществляется. Например, чтобы построить систему информационной поддержки научной и производственной деятельности в определенной области знаний (ОЗ), необходима формализация не только функциональности такой системы, но и семантики назначенной ей области знаний.

Наиболее адекватным и эффективным средством формализации областей знаний являются онтологий [Guarino, 1998]. В настоящее время формализм онтологий является своего рода универсальным семантическим кодом (УСК) [Мартынов, 1974], так как он повсеместно используется для построения формальных моделей областей знаний с целью фиксации общего разделяемого всеми экспертами (или, по крайней мере, их большинством) знания об этих областях [Gruber, 1995].

Формализация семантики области знаний в виде онтологий служит не только целям компактного и непротиворечивого представления моделируемой области знаний, она также формирует понятийный базис для представления тех знаний о данной ОЗ,

которые невозможно, неудобно или неэффективно представлять средствами онтологий. В частности, в ее терминах описывается семантика данных и информационных ресурсов, используемых в информационной системе, а также экспертные правила и другие компоненты базы знаний интеллектуальной системы. Онтология также может служить основой для других моделей представления знаний, как например, для интегрированной модели представления знаний [Загорулько, 2013], объединяющей различные взаимодополняющие друг друга методы и средства представления и обработки знаний и служащей базисом для оригинальной технологии разработки интеллектуальных систем [Загорулько, 2013a].

Таким образом, онтология занимает центральное место в системе знаний любой информационной системы, в том числе интеллектуальной. Можно также сказать, что система знаний любой интеллектуальной системы символического (семиотического) типа базируется на онтологий, только она не всегда явно задана.

В докладе обсуждается роль онтологий в интеллектуальных и информационных системах, рассматриваются средства формализации семантики области знаний на основе онтологий, приводятся и обсуждаются примеры практического использования онтологий в интеллектуальных и информационных системах, в разработке которых принимали участие авторы доклада.

1. Роль онтологии в интеллектуальных и информационных системах

Онтология выступает в качестве базового средства формализации областей знаний для широкого класса интеллектуальных и информационных систем, в частности, для экспертных систем, систем поддержки принятия решений и различных систем информационной поддержки научной и производственной деятельности.

Экспертные системы [Попов, 1987] относятся к классу программных систем, аккумулирующих знания специалистов в конкретных предметных областях и тиражирующих этот эмпирический опыт для консультаций менее квалифицированных пользователей. В связи с этим основным и наиболее важным компонентом ЭС является база знаний. Именно полнота и непротиворечивость представленных в ней знаний, определяют мощность экспертной системы и качество получаемых ею решений.

Онтология может составлять каркас базы знаний ЭС, создавать базис для описания основных понятий предметной области (ПО) и служить основой для интеграции баз данных, содержащих фактические знания, необходимые для полноценного функционирования экспертной системы. Кроме того, в терминах онтологии могут описываться экспертные правила и прецеденты [Варшавский и др., 2006], что значительно повышает их уровень описания и «понимаемость» пользователями-экспертами.

Интеллектуальные системы поддержки принятия решений (ИСППР) [Петровский, 2009] помогают лицу, принимающему решения (ЛПР), использовать данные и модели для решения его профессиональных слабо формализуемых задач. ИСППР и ЭС относятся практически к одному классу интеллектуальных систем. Очень часто в ИСППР включается несколько различных ЭС, поэтому все, что было сказано выше о роли онтологии в ЭС, относится и к ИСППР. В то же время имеются аспекты использования онтологий, специфичные для СППР.

Так, из-за того, что ИСППР в основном предназначена для решения плохо формализуемых задач, очень важно иметь детальное непротиворечивое описание проблемной области, в рамках которой система осуществляет поддержку решения задач ЛПР. И онтология является незаменимым инструментом для создания такого описания.

В большинстве ИСППР используются большие массивы разнородных данных и знаний. Благодаря тому, что онтология позволяет явно описывать семантику данных и знаний, она обеспечивает базис для их интеграции и совместного использования при решении задач.

Система информационной поддержки научной и

производственной деятельности [Ануреев и др., 2009], как правило, должна обеспечивать своим пользователям представление всей информации об определенной области знаний и/или производства, о ее составляющих (разделы/подразделы науки и/или отрасли/подотрасли производства, объекты, методы и техники исследования и т.п.), а также о субъектах (участниках) научной и/или производственной деятельности (персоналий, групп, сообществ и других организаций, включенных в процесс исследования и/или производства). При этом она должна предоставлять содержательный доступ ко всем перечисленным знаниям и данным, т.е. удобную навигацию по всему информационному пространству системы и возможность поиска информации в терминах представляемой области знаний.

В системах такого класса онтология задает формальное описание области знаний, на основе которого строится информационное наполнение (контент) системы, выполняется интеграция в единое информационное пространство (ИП) релевантных информационных ресурсов и документов. На основе онтологии строится пользовательский интерфейс, обеспечивающий содержательный доступ к знаниям и данным, интегрированным в информационное пространство системы. В частности, в таком интерфейсе пользователь может использовать онтологию в качестве «проводника» для навигации по ИП системы, а также формулировать поисковые запросы, основными элементами которых являются понятия и отношения онтологии.

Заметим также, что при разработке интеллектуальных и информационных систем возможно использование онтологий, ранее разработанных для области знаний этих систем. Это позволяет переиспользовать уже проверенные на практике знания, что обеспечивает высокое качество создаваемых систем и их потенциальную интегрируемость с ранее разработанными.

Таким образом, онтологии могут использоваться как при разработке интеллектуальных и информационных систем, так и в качестве полноправного компонента в процессе их функционирования.

2. Средства формализации знаний на основе онтологий

Существует множество подходов к формализации знаний в интеллектуальных и информационных системах на основе онтологий. В этой главе будут рассмотрены наиболее популярные из них, а также подходы, разработанные в нашем коллективе.

2.1. Средства формализации знаний в программной среде SemP-TAO

Рассмотрим подход к формализации знаний, предлагаемый в программной среде SemP-TAO

[Загорулько Ю.А. и др., 1996], предназначенной для построения широкого класса интеллектуальных систем.

Модель представления знаний программной среды SemP-ТАО объединяет на основе объектно-ориентированного подхода такие классические средства представления и обработки знаний, как семантические сети и продукционные правила, а также недоопределенные модели [Нариньяни и др., 1998] и методы программирования в ограничениях. Благодаря такому сочетанию средств, в данной среде эффективно решаются задачи, которые требуют сочетания логического вывода и вычислений над неточно (частично) заданными значениями.

Язык представления и обработки знаний (ЯПОЗ) технологической среды SemP-ТАО включает два типа средств – средства для описания предметной области и средства для спецификации приложений, предназначенных для использования в рамках описанной предметной области.

Первая группа средств позволяет представлять предметную область в виде классов объектов и отношений, а также конструировать необходимые для этого новые типы данных и домены. Описания классов и отношений составляют онтологию, которая задает семантику и структуру семантической сети – основного средства представления декларативных знаний в SemP-ТАО.

Вторая группа средств позволяет описывать функциональность приложения, главным образом, в виде системы продукционных правил, работающей над семантической сетью. Язык содержит не только продукционные правила и средства оперирования семантической сетью, включая операторы создания объектов и бинарных отношений, их редактирования (изменения значений атрибутов объектов) и удаления из сети, но и средства динамического управления активацией продукционных правил, а также. Кроме того, язык включает традиционные средства императивного программирования, а также средства для создания эргономичных пользовательских интерфейсов.

Условие применения правила состоит из образца, включающего описание объектов и отношений, заданных в онтологии, а также глобальных и локальных предикатов, накладывающих ограничения, соответственно, на контекст, в котором применяется данное правило, и на значения аргументов отношений и атрибутов объектов, описанных в образце.

В правой части правила разрешается обращаться к операторам, оперирующим семантической сетью, а также другим операторам и функциям, разрешенным в языке. В частности, включение в правую часть правил средств динамического управления активацией правил позволяет обеспечить высокую гибкость управления процессом вывода и обработки информации.

Возможность использования в посылках и заключениях продукционных правил понятий (классов) и отношений, определенных в онтологии, значительно повышает уровень описания правил и их «понимаемость» экспертами.

2.2. Формализация знаний на основе дескриптивных логик

Дескриптивные логики (description logics) – это семейство логик, специально спроектированных для представления знаний о предметной области [Baader et al., 2003; Тузовский и др., 2005].

Дескриптивные логики (ДЛ) должны были прийти на смену таким средствам и формализмам представления знаний, как семантические сети, фреймы и исчисление предикатов, объединив их достоинства и избавившись от их недостатков. Известно, что семантические сети и фрейм-представления, обладая высокими по выразительности средствами для представления знаний о предметной области, не обеспечивают достаточно формального описания ПО. В то же время исчисление предикатов дает высокий уровень формализации описаний, но не предоставляет удобных средств для представления специфических знаний о предметной области.

Для того, чтобы ДЛ можно было использовать в качестве формализма представления знаний в интеллектуальных системах, необходимо, чтобы такие системы отвечали на запросы пользователей за разумное время. В связи с этим процедуры логического вывода, применяемые в ДЛ, в отличие, например, от алгоритмов доказательства теорем логик первого порядка, должны всегда завершаться, как при положительных, так и при отрицательных ответах. Поэтому ДЛ часто рассматривается как область разрешимых частей логики первого порядка.

Разрешимость и сложность задач вывода зависит от выразительной мощности конкретной дескриптивной логики, построенной в рамках концепции ДЛ. С одной стороны, очень экспрессивные ДЛ сталкиваются с задачами вывода высокой сложности, которые могут быть даже неразрешимыми. С другой стороны, очень слабые ДЛ (с эффективными процедурами вывода) могут не обладать достаточной выразительностью для представления важных понятий конкретной прикладной задачи. Определение баланса между выразительностью конкретной дескриптивной логики и сложностью задач вывода является одной из наиболее важных проблем при ее конструировании (или выборе).

Базовыми элементами ДЛ являются концепты (понятия) и роли. Каждый *концепт* представляет класс, категорию или сущность. *Роли* служат для описания свойств концептов и отношений между ними. В ДЛ также входят *конструкторы (операции)*, с помощью которых можно строить из

базовых элементов понятия и роли более высокого уровня.

База знаний, построенная средствами дескриптивной логики представляется в виде пары *Tbox* и *Abox*. *Tbox* (*terminological knowledge*) – это набор утверждений, описывающих терминологический словарь базы знаний, т.е. набор классов, их свойства и отношения между ними (интенциональные знания). *Abox* (*assertional knowledge*) представляет собой реализацию схемы классов в виде набора экземпляров, содержащих утверждения об экземплярах понятий (экстенциональные знания). По существу *Tbox* является моделью того, что *может* быть истинным, а *Abox* является моделью того, что *в настоящее время* является истинным.

Дескриптивный язык ДЛ имеет модельно-теоретическую семантику [Tarski, 1956]. Поэтому утверждения в *TBox* и в *ABox* могут отождествляться с формулами в ЛПП или, в некоторых случаях, незначительными их расширениями.

Благодаря тому, что ДЛ имеют формальную семантику, они позволяют не только описать терминологию и утверждения, но также предоставляют возможности по выполнению на их основе логического вывода (автоматического доказательства). Автоматическое доказательство обеспечивает ответы на такие запросы, как: выполнимость понятий (*satisfiability*) – может ли существовать некоторое понятие *C*, описание которого задано в *Tbox*; включение (*subsumption*) – является ли некоторое понятие *C* частным случаем (подмножеством) понятия *D*; согласованность (*consistency*) – является ли вся БЗ согласованной; проверка экземпляра (*instance checking*) – является ли некоторое утверждение истинным.

Понятия могут строиться из выражений. В состав операторов включается базовый набор операций (дополнение, объединение и пересечение множеств) и количественные ролевые ограничения.

С помощью автоматического доказательства можно формировать ответы на запросы к *Tbox* и *Abox*. Запросы к *Tbox* определяют классификацию и отношения между понятиями, а запросы к *Abox* определяют текущее состояние известных фактов. Выполнимость понятия является доказательством того, что понятие, или понятийное выражение является логически согласованным с БЗ. Кроме этого операция включения ДЛ (*subsumption*) может использоваться для обнаружения отношений класс-подкласс, которые заданы в БЗ неявно.

Таким образом, ДЛ определяют формальный язык для описания семантики ПО вместе с теорией доказательства (процедурами логического вывода).

Достоинством ДЛ является то, что они позволяют строить формальные декларативные языки, обладающие большой выразительной мощностью для представления знаний. Другое их

достоинство – применяемые в них процедуры логического вывода, как правило, завершаются, как при положительных, так и при отрицательных ответах. В то же время ДЛ трудны для освоения экспертами, которые не являются специалистами в области математики и программирования.

2.3. Средства формализации знаний подхода Semantic Web

Средства подхода Semantic WEB – это набор методов и технологий, предназначенных для представления информации в пригодном для машинной обработки виде. Эти средства позволяют представлять информацию в виде семантической сети, специфицированной с помощью онтологий. Благодаря стандартизованному представлению программа-клиент может непосредственно извлекать факты и делать из них логические заключения, используя протокол HTTP и идентификаторы ресурсов URI. Использование средств Semantic WEB стало в последние 10 лет очень популярным для разработки интеллектуальных приложений.

Наиболее выразительным средством описания онтологий для семантической паутины, является язык OWL (*Web Ontology Language*) [OWL, 2013]. В основе этого языка – представление действительности в модели данных – «объект – свойство». OWL пригоден для описания не только веб-страниц, но и любых объектов действительности. Каждому элементу описания в этом языке (в том числе свойствам, связывающим объекты) ставится в соответствие URI.

OWL обеспечивает лучшую выразительность и более богатую семантику, чем многие другие языки описания структур, такие как XML, RDF, RDF Schema (RDF-S) [Лапшин, 2010] и другие.

Язык OWL имеет три диалекта (в порядке возрастания выразительности):

- OWL Lite является самым простым диалектом и поддерживает только классификационную иерархию понятий, элементарные ограничения и простейшие виды аксиом.
- OWL DL обеспечивает максимальную выразительность при сохранении полноты вычислений (все заключения гарантировано будут вычислимыми) и разрешимости (все вычисления производятся за конечное время). OWL DL включает все языковые конструкции OWL, но ограничивает свободу их использования (например, в то время как класс может быть подклассом многих классов, класс не может быть представителем другого класса). OWL DL так назван из-за его соответствия дескриптивным логикам (в частности, SHOIN(D)).
- OWL Full предоставляет максимальную выразительность и синтаксическую свободу RDF, но без гарантий вычислений за конечное время. Например, в OWL Full класс может рассматриваться

одновременно как собрание индивидов и как один индивид в своем собственном значении.

При использовании OWL для построения онтологии интеллектуальных или информационных систем наиболее часто используют диалекта OWL DL, поскольку имеющиеся машины вывода рассчитаны на поддержку OWL только в рамках данного диалекта.

Онтологии, написанные на языке OWL, включают в себя описание классов и их свойств, а также экземпляры классов (индивидов), совокупность которых вместе со связывающими их отношениями образует семантическую сеть.

Для записи OWL-онтологий могут использоваться разные форматы: RDF/XML, OWL/XML, Манчестерский синтаксис.

Resource Description Framework (далее – RDF) – это разработанная консорциумом W3C модель для представления данных, в особенности – метаданных. RDF представляет утверждения о ресурсах в виде, пригодном для машинной обработки.

Для поиска информации в онтологии используются такие языки как SPARQL и DL Query.

SPARQL (рекурсивный акроним от английского SPARQL Protocol and RDF Query Language) — это язык запросов к данным, представленным по модели RDF, а также протокол для передачи этих запросов и ответов на них. SPARQL является рекомендацией консорциума W3C и одной из технологий Semantic WEB. Существуют реализации языка запросов SPARQL для ряда языков программирования.

Язык DL Query предназначен для осуществления запросов к онтологии и основан на Манчестерском синтаксисе.

Логический вывод в OWL-онтологиях осуществляется на основе аксиом онтологии средствами одной из машин вывода, каковых уже разработано достаточное количество (Pellet, FaCT++, HermiT). Помимо аксиом можно задавать правила вывода, которые позволяют выводить информацию, не содержащуюся явным образом в онтологии. Для этого можно использовать язык SWRL (Semantic Web Rule Language), основанный на объединении языков OWL-DL и RuleML. Ядром RuleML является язык Datalog, который, в свою очередь, является синтаксическим подмножеством языка Prolog. Этим и объясняется то, что SWRL-правила состоят из хорновских дизъюнктов.

Для работы с OWL онтологиями в программных приложениях используются различные библиотеки. Наиболее популярными из них являются OWL API и Jena.

Для разработки онтологий создано множество редакторов. Наиболее популярным является Protégé [Protégé, 2013]. Protégé – это свободно распространяемый редактор онтологий с открытым

кодом и одновременно фреймворк для построения баз знаний. Онтологии, построенные в Protégé, могут быть экспортированы во множество форматов, включая RDF (RDF Schema), OWL, и XML Schema.

Protégé имеет открытую, легко расширяемую архитектуру за счет поддержки модулей расширения функциональности. Это позволяет настраивать редактор под нужды различных приложений. Популярность Protégé объясняется и тем, что он поддерживается значительным сообществом, включающим разработчиков и ученых, правительственных и корпоративных пользователей.

Таким образом, подход Semantic Web дает богатый набор средств формализации знаний, который постоянно расширяется и развивается благодаря поддержке консорциума W3C. Эти средства обладают не только значительной выразительностью, но еще и являются разрешимыми. Пока слабым звеном этим средств является язык SWRL, который позволяет описывать очень простые правила, выразительности которых недостаточно для разработки даже примитивных экспертных систем.

Другой проблемой в использовании средств, предоставляемых подходом Semantic Web, является их высокая сложность для простых пользователей, не знакомых с современными логиками. Это затрудняет использование этих средств в технологиях построения информационных и интеллектуальных систем, ориентированных на экспертов.

2.4. Средства формализации знаний в технологии построения порталов научных знаний

Рассмотрим подход к формализации семантики области знаний на основе онтологий, который использован в технологии построения порталов научных знаний [Загорюлько и др., 2008].

Эта технология предназначена для разработки порталов научных знаний (ПНЗ) – специализированных интернет-порталов, обеспечивающих интеграцию и систематизацию научных знаний и информационных ресурсов определенной области знаний, а также содержательный доступ к ним из любой «точки» интернет-пространства. Портал научных знаний не только предоставляет доступ к собственным информационным ресурсам, но и поддерживать навигацию по заранее размеченным (прондексированным) ресурсам, размещенным в сети Интернет.

Технология ориентирована на создание и сопровождение порталов экспертами, не являющимися специалистами в области информатики и программирования, и с минимальным участием (на первых этапах разработки) инженеров знаний. Важной особенностью технологии является обеспечение

возможности декларативной настройки портала на заданную область знаний не только в процессе создания, но и в ходе эксплуатации. Реализация такой возможности позволяет отслеживать динамику появления новых знаний и информационных ресурсов по тематике портала и, тем самым, обеспечит поддержку его актуальности и полезности.

Создание технологии с такими свойствами стало возможным благодаря выбору онтологии в качестве средства формализации семантики области знаний и использованию ее в качестве концептуальной основы и информационной модели портала знаний онтологии.

Для того, чтобы онтологии успешно выполняли функцию формализации семантики моделируемых областей знаний, требуются удобные и эффективные средства их представления, а также методологии (технологии), облегчающие процесс их построения и делающие их более надежными.

Прежде всего, для представления онтологии необходим формализм, обеспечивающий описание понятий моделируемой области знаний, их свойств и разнообразных семантических связей (отношений) между ними. Важным требованием к формализму описания онтологии является возможность выстраивания понятий в иерархию «общее–частное» и поддержка наследования свойств по этой иерархии. Этот формализм должен также предоставлять возможность задания ограничений на значения возможных свойств (атрибутов) объектов – экземпляров понятий онтологии.

Для повышения компактности описания области знаний, а также для поддержки дополнительного контроля за использованием отношений, формализм должен предоставлять возможность приписывания отношениям различных математических свойств, таких как симметричность, транзитивность, антисимметричность и др

Кроме того, для представления сложно-структурированной информации требуются бинарные отношения, снабженные дополнительными атрибутами, специализирующими связь между двумя сущностями (аргументами отношения). Примером такого отношения может быть, например, отношение, связывающее между собой персону и организацию, в которой эта персона работает. Дополнительные атрибуты здесь нужны для того, чтобы отразить, в какой должности работает данная персона в этой организации и с какого времени.

Такой формализм был разработан (его подробное описание можно найти в [Загоруйко, 2013а]) и используется в качестве базиса при построении онтологий порталов научных знаний.

На рисунке 1 представлены основные компоненты данной технологии. Она включает в себя набор универсальных и базовых онтологий,

методологию построения онтологий и программные средства поддержки процесса построения и редактирования онтологий и тезауруса, заполнения и редактирования контента портала, средств поиска и навигации по онтологии и контенту, а также пользовательский интерфейс.

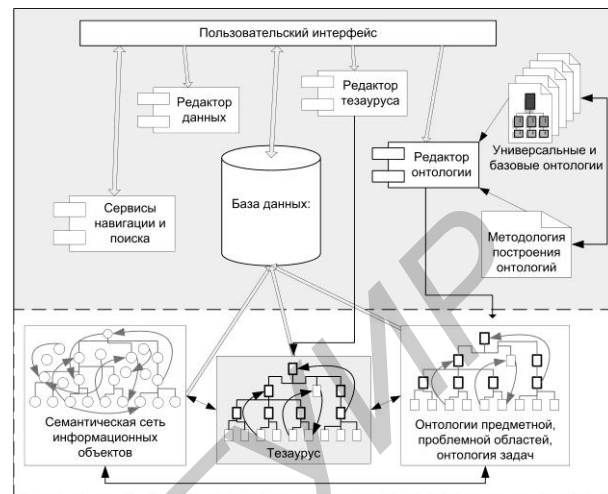


Рисунок 1 – Технология построения порталов научных знаний

Для обеспечения унифицированного представления разнородных знаний и данных, учета их связанности, а также поддержки функциональности ПНЗ предложена информационная модель, которая объединяет модели проблемной области и области (областей) знаний портала, а также построенную на их основе модель представления данных.

В качестве ядра, базового компонента информационной модели портала научных знаний была выбрана онтология, что позволило не только обеспечить его основную функциональность, но и сделало настраиваемым на любую область научных знаний.

Создание портала начинается с построения его онтологии. Согласно предложенной методологии эта онтология создается на основе базовых онтологий путем специализации и расширения имеющейся в них системы понятий. Это значительно облегчает создание онтологии и ее дальнейшее сопровождение.

В качестве базовых выбраны онтология научной деятельности и онтология научного знания, на основе которой строится онтология области знаний портала.

Онтология научной деятельности, выступающая в качестве онтологии проблемной области портала знаний, базируется на онтологии, предложенной в [Benjamins et al., 1998] для описания научно-исследовательских проектов и расширенной в [Загоруйко и др., 2008] для применения к более широкому классу задач. Она включает базовые классы понятий, относящиеся к организации научной и исследовательской деятельности, такие как *Исследователь*, *Организация*, *Событие*, *Деятельность (Проект)*, *Публикация* и др. В эту

онтологию также включен класс *Информационный ресурс*, который служит для описания информационных ресурсов, представленных в сети Интернет.

Онтология научного знания фиксирует основные содержательные структуры, используемые для построения предметных онтологий. В частности, эта онтология содержит мета-понятия, задающие структуры для описания понятий конкретной области знаний, такие как *Раздел науки*, *Метод исследования*, *Объект исследования*, *Научный результат*.

Понятия базовых онтологий связаны между собой ассоциативными отношениями, выбор которых осуществляется не только исходя из полноты представления проблемной и предметной областей портала, но и с учетом удобства навигации по его контенту и поиска информации.

Понятия онтологии области знаний портала являются реализациями метапонятий онтологии научного знания и могут быть упорядочены в иерархию «общее-частное».

Вводя формальные описания понятий предметной области в виде классов объектов и отношений между ними, онтология портала задает структуры для представления реальных объектов и связей между ними. В соответствии с этим данные на портале представлены в виде семантической сети, т.е. как множество разнотипных информационных объектов и связей между ними, которые в совокупности образуют информационное содержание (контент) портала.

Для управления онтологиями используется редактор онтологий, реализованный как web-приложение и доступный через Интернет. Этот редактор проектировался и разрабатывался таким образом, чтобы он был понятен экспертам, прост и удобен в использовании. В частности, из-за этих требований мы отказались от такого популярного средства построения онтологий как редактор Protégé [Protégé, 2013].

Ввод информации осуществляется с помощью управляемого онтологией редактора данных, который позволяет создавать, редактировать и удалять информационные объекты (ИО) и связи между ними. Формы для ввода конкретных ИО и их связей автоматически генерируются по онтологии.

Содержательный доступ к систематизированным знаниям и информационным ресурсам заданной области знаний обеспечивается с помощью предоставляемых порталом развитых средств навигации и поиска, функционирование которых также базируется на онтологии. Благодаря этому основной сценарий работы пользователя с порталом состоит из выбора либо с помощью средств визуализации онтологии, либо с помощью механизма поиска объектов определенного класса, их просмотра, навигации по их связям (реализациям

отношений онтологии) и фильтрации списков таких объектов.

Таким образом, данная технология поддерживает разработку порталов научных знаний без участия разработчиков-программистов. Для создания портала требуются только инженеры знаний – специалисты в представлении знаний и эксперты – носители знаний в моделируемой области. Использование онтологий позволяет сделать процесс разработки портала научных знаний более технологичным, в частности, переиспользовать ранее полученные и формализованные знания.

2.5. Использование онтологий в технологии разработки тематических интеллектуальных научных интернет-ресурсов

Подход к построению порталов научных знаний получил развитие в технологии разработки тематических интеллектуальных научных интернет-ресурсов (ИНИР) [Загорулько и др., 2013b], которые, как и ПНЗ, предназначены для осуществления информационной поддержки научной и производственной деятельности в определенной области знаний. Основу ИНИР, как и порталов научных знаний составляют онтологии.

Главное отличие данной технологии от технологии ПНЗ состоит в существенном использовании семантических web-сервисов. Такие сервисы в технологии ИНИР выступают и в качестве внутренних функциональных компонентов ИНИР, расширяющих функциональность ИНИР, и в качестве внешних информационных ресурсов, накапливаемых и систематизируемых в ИНИР с тем, чтобы обеспечить к ним содержательный доступ.

Тематический ИНИР представляет собой доступную через Интернет информационную систему, обеспечивающую систематизацию и интеграцию научных знаний и информационных ресурсов определенной области знаний, содержательный эффективный доступ к ним (поиск и навигацию) и поддерживающую их использование при решении различных научных и производственных задач за счет предоставления соответствующих интерфейсов и сервисов.

Система знаний ИНИР базируется на формализмах онтологий и семантических сетей. При этом онтология составляет ядро системы знаний и наряду с описанием моделируемой области знаний содержит соотношенное с ним описание структуры и типологии интегрируемых информационных ресурсов и методов интеллектуальной обработки содержащихся в них данных и знаний.

Семантическая сеть, структура которой определяется онтологией ИНИР, играет роль интеллектуального хранилища данных, в котором накапливается информация о релевантных научных информационных ресурсах и web-сервисах,

реализующих методы обработки содержащейся в них информации.

На основе онтологии и семантической сети организуется удобная навигация по научным знаниям и информационным ресурсам, интегрированным в ИНИР, а также содержательный поиск требующихся данных и средств их интеллектуальной обработки (представленных, в том числе, в виде web-сервисов).

Особенностью данного подхода является то, что ИНИР позволяет значительно сократить время, которое требуется для обеспечения доступа к необходимой информации и ее анализа, за счет предварительного поиска релевантных интернет-ресурсов и аккумуляции их описаний непосредственно в семантической сети ИНИР. С этой целью в программную оболочку ИНИР включается подсистема сбора онтологической информации (метаданных) о релевантных интернет-ресурсах, которая в своей работе опирается на онтологию и тезаурус области знаний ИНИР.

Система знаний ИНИР (см. рисунок 2) включает онтологию ИНИР, тезаурус области знаний ИНИР, семантическую сеть, служащую для представления информационного содержания (контента) ИНИР, а также интегрируемые в ИНИР информационные ресурсы и средства их интеллектуальной обработки (web-сервисы).



Рисунок 2 – Система знаний ИНИР

Онтология ИНИР состоит из трех взаимосвязанных онтологий, отвечающих за представление указанных выше компонентов знаний, а именно: онтологии области знаний ИНИР, онтологии научных интернет-ресурсов и онтологии задач и методов.

Онтология области знаний ИНИР строится на основе двух базовых онтологий, включенных в программную оболочку ИНИР, – онтологий научной деятельности и онтологий научного знания, унаследованных из технологии построения ПНЗ.

Основным классом онтологии научных интернет-ресурсов является класс *Информационный ресурс*, который предназначен для описания, представленных в сети Интернет информационных ресурсов. Набор атрибутов и связей этого класса основан на стандарте Dublin core [Hillmann., 1995-2013]. Его атрибутами являются: название ресурса, язык ресурса, тематика ресурса, тип доступа к ресурсу и т.п. Объекты этого класса могут быть связаны семантическими отношениями с другими информационными объектами, представляющими в контенте ИНИР организации, персоны, публикации, события, разделы науки и т.д.

Онтология задач и методов кроме описания задач, на решение которых нацелен ИНИР, и методов их решения включает также описания web-сервисов, реализующих методы обработки информации, содержащейся в интегрируемых в ИНИР информационных ресурсах. Описания таких web-сервисов базируются на онтологии OWL-S [OWL-S, 2004], предназначенной для описания семантических web-сервисов (Semantic Web Services) [Cabral, 2004] и представленной на языке OWL.

Таким образом, онтология ИНИР, вводя формальные описания понятий некоторой области знаний, типов интегрируемых информационных ресурсов и методов их интеллектуальной обработки в виде классов объектов и отношений между ними, одновременно задает структуры для представления информации о реальных объектах моделируемой области знаний, интегрируемых информационных ресурсах и методах и средствах обработки содержащейся в них информации. Данная информация хранится в контенте ИНИР в виде объектно-ориентированной семантической сети, типы информационных объектов и отношений которой определяются классами объектов и отношений, введенных в онтологии ИНИР.

Тезаурус в ИНИР служит для описания смысла понятий (терминов), используемых в моделируемой области знаний. При этом тезаурус позволяет задавать смысл понятий не только с помощью определений, но и посредством соотнесения их с другими понятиями. Благодаря этому тезаурус может применяться как при обработке пользовательских запросов, так и при поиске и аннотировании информационных ресурсов [Лукашевич, 2011], интегрируемых в ИНИР.

Тезаурус ИНИР строится на основе ядра тезауруса, изначально включенного в систему знаний ИНИР. Ядро тезауруса содержит описание понятий базовых онтологий, включая описание терминов, с помощью которых они представляются в интернет-ресурсах.

ИНИР имеет традиционную трехуровневую архитектуру, включающую уровень доступа к информации, уровень обработки информации и базовый уровень.

Первый уровень обеспечивается пользовательским интерфейсом – программным компонентом, предоставляющим конечному пользователю содержательный доступ к контенту ИНИР при решении его задач.

Уровень обработки информации обеспечивает все информационные потоки в ИНИР – от конструирования онтологии до обработки пользовательских запросов. Он включает модуль поиска информации в информационном пространстве ИНИР, средства разработки/настройки базы знаний ИНИР и управления его контентом, а также подсистему сбора онтологической информации (метаданных) об интернет-ресурсах.

Базовый уровень обеспечивает управление всеми данными и знаниями информационного пространства ИНИР. Он выполняет функции хранения и манипулирования данными (контентом ИНИР) и знаниями (онтологией и тезаурусом) с использованием средств стандартных СУБД (MySQL), технологий Semantic Web (OWL, RDF) и семантических web-сервисов (WSDL, OWL-S). В частности, в качестве хранилища данных используется свободно распространяемая версия системы Virtuoso (Virtuoso Open-Source Edition) [Virtuoso, 2013], обеспечивающая эффективную работу с большими объемами данных в RDF-формате (RDF Triple Store) [Erling, 2009]. Кроме того, эта система предоставляет возможность взаимодействия с web-сервисами.

3. Использование онтологий в ЭС и СППР

Как было показано выше, онтологии используются как при разработке ЭС и СППР, так и во время их функционирования.

3.1. Использование онтологий в медицинских ЭС. Экспертная система для подбора оздоровительной программы

Данная система предназначена для оказания помощи человеку при подборе подходящей для него оздоровительной программы.

На основе данных о конкретном человеке (антропометрические данные, список заболеваний и проблем, связанных с его здоровьем и самочувствием, преследуемые им цели) система предлагает ему наиболее подходящие виды физических нагрузок, диеты, восстановительные или профилактические процедуры. Кроме того, система может выдать справочную информацию о восстановительных комплексах, их показаниях и противопоказаниях.

Знания в системе представлены OWL-онтологией, фрагмент которой показан на рисунке 3.

OWL-онтология создана средствами редактора Protégé. Логический вывод в рассматриваемой системе реализуется на основе заданных в онтологии аксиом средствами машины вывода

Pellet. Аксиомы задают ограничения для понятий онтологии. Кроме того, в системе используются запросы на языке SPARQL, которые также записываются в терминах онтологии.

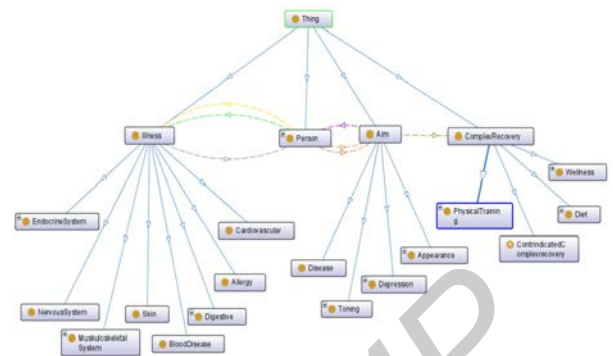


Рисунок 3 – Онтология комплексного восстановления организма человека.

3.2. Использование онтологий в системе поддержки принятия решений на нефтегазодобывающем предприятии

Покажем, как используются онтологии в разработанной нами СППР [Загорюлько и др., 2010], которая на основе анализа статической и динамической информации об объектах технологической инфраструктуры нефтегазодобывающего предприятия должна выработать для ЛПР рекомендации по предотвращению аварийных ситуаций, улучшению показателей работы предприятия, о проведении планового технического обслуживания и/или экстренного ремонта объектов, о списании объектов и замене их новыми и пр.

Данная СППР имеет достаточно гибкую архитектуру, позволяющую подключать к ней различные методы решения задач и решатели. Эта гибкость во многом достигается благодаря тому, что в систему в качестве полноправного компонента включены две взаимосвязанные онтологии – онтология предметной области и онтология задач, обеспечивающие настройку системы на предметную область и типы решаемых задач.

Базовыми понятиями онтологии ПО (см. рисунок 4) являются: *Объект*, *Нормативно-справочный объект*, *Состояние*, *Результат*.

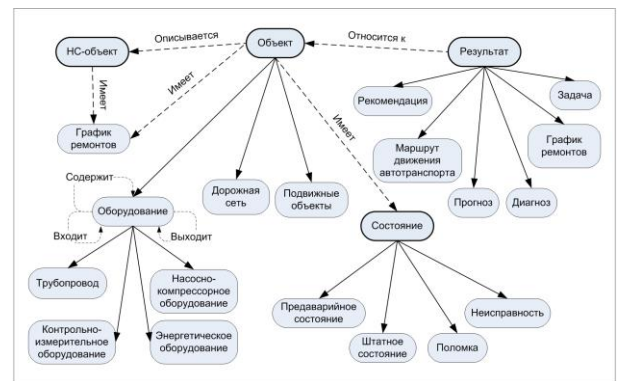


Рисунок 4 – Фрагмент онтологии ПО.

В классе *Объект* выделяются подклассы *Оборудование* (насосы, трансформаторы, трубопроводы и т.п.), *Подвижные объекты* (автоцистерны, грузовики, автобусы и другие виды автотранспорта) и *Дорожная сеть*. Объекты могут находиться в том или ином состоянии, для описания которого вводится класс *Состояние*.

Важное место в онтологии ПО занимает описание топологии технологической инфраструктуры предприятия: взаимное расположение и связность объектов задается отношениями «Содержит», «Входит», «Выходит».

Для контроля соответствия параметров объектов анализа нормативным значениям используется нормативно-справочная информация: для каждого понятия, представляющего некоторый вид оборудования или контейнера оборудования, заводится специальное понятие, которое мы будем называть нормативно-справочный объект или *НС-объект*. Этот *НС-объект* имеет тот же набор атрибутов, что и исходное понятие (объект), но числовые значения его атрибутов задаются парой чисел, которые определяют интервал изменения атрибута данного понятия согласно нормативу (техпаспорту).

Результатами работы СППР являются *Диагнозы* состояния оборудования и подвижных объектов, *Рекомендации* для ЛПР, *Прогнозы* изменения состояния объектов, решения транспортных задач (*Маршруты движения автотранспорта*), скорректированные *Графики ремонтов* и т.п. Отдельным результатом может быть порожденная исполняемым модулем поддержки принятия решений *Задача*, в которой, например, более детально анализируется тот или иной объект.

В данной системе онтология ПО, с одной стороны, выступает в качестве высокоуровневого интерфейса к внутреннему хранилищу данных, обеспечивая доступ к ним в виде объектов предметной области, с другой стороны, она определяет формат представления данных в самой СППР в виде тех же объектов предметной области (экземпляров понятий онтологии) и отношений между ними, что позволяет упростить и унифицировать обмен информацией между разнородными компонентами и модулями (решателями) СППР. В терминах онтологии также описываются экспертные правила (см. ниже).

Онтология задач (см. рисунок 5), описывая типы решаемых системой задач, определяет тем самым ее функциональность. Она включает описания задач, модулей, реализующих их решения, и используемых ими решателей.

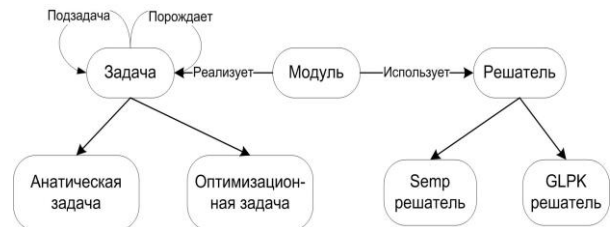


Рисунок 5 – Фрагмент онтологии задач.

Для решения аналитических задач используются модули поддержки принятия решений, реализованные в парадигме продукционной модели. При этом для спецификации входящих в эти модули продукционных правил, а также обеих онтологий используется язык представления и обработки знаний системы Semp-TAO, что позволяет использовать введенные в онтологиях понятия и отношения в посылках и заключениях продукционных правил. Для исполнения этих модулей поддержки принятия решений используется интерпретатор продукционных правил системы Semp-TAO (Semp-решатель).

Приведем пример продукционного правила:

FORALL NR: P: *Насос* (*Capacity*, *Energy_consumption*, *Efficiency*), *НС_Насос* (*Efficiency_min*),
Описывается (P, NR),
Efficiency > *Efficiency_min* &
Tendency (*Capacity*) = “*падаем*” &
Tendency (*Efficiency*) = “*падаем*” &
Tendency (*Energy_consumption*) = “*растем*”
=>
Create Прогноз (*message*: “*КПД достигнет критического уровня через*” + *Critical_moment* (*Efficiency*, *Efficiency_min*) + “*часов*”, *parameters*: {<Pump, P>});
Create Задача (*name*: “*Диагностика насоса*”, *parameters*: {<Pump, P>}).

Заметим, что в этом правиле *Capacity*, *Energy_consumption* и *Efficiency* являются временными рядами параметров насоса *подача*, *потребляемая энергия* и *кпд насоса P*, а NR – это нормативно-справочный объект насоса P.

Это правило говорит о том, что если для какого-либо насоса возрастает потребление электроэнергии при одновременном снижении кпд и подачи, и эти показатели, оставаясь в нормативных рамках, приближаются к их границам, то с определенной долей уверенности можно прогнозировать, что этот насос скоро выйдет из строя. При возникновении такой ситуации это правило выдаст соответствующее сообщение ЛПР и запустит задачу детальной диагностики проблемного насоса.

4. Использование онтологий в системах информационной поддержки научной и производственной деятельности

4.1. Использование онтологий в порталах научных знаний

Рассмотренная в разделе 2.4 технология построения порталов научных знаний была использована в рамках проектов по созданию научных интернет-порталов по компьютерной лингвистике и археологии, выполнявшихся на базе ИСИ СО РАН.

Разработанный совместно с Институтом археологии и этнографии СО РАН **археологический портал знаний** [Андреева и др., 2006] был создан для решения задачи интеграции накопленных знаний и информационных ресурсов по археологии, а также обеспечения содержательного доступа к ним. Этот портал рассчитан на широкий круг пользователей – от научных работников и преподавателей до студентов и школьников, интересующихся достижениями археологической науки.

Онтология области знаний портала включает четыре базовых иерархии:

- иерархию *Разделов науки*, отражающую иерархию основных направлений научной деятельности в археологии. Примерами таких направлений являются *Общая*, *Полевая* и *Реконструктивная археология*, в свою очередь подразделяющиеся на более частные;
- иерархию *Объектов*, определяющую объекты исследования в археологии, примерами которых могут являться археологические культуры, этносы, исторические личности, предметы быта, обнаруженные в процессе раскопок;
- иерархию *Методов исследования*, служащую для описания подходов, принципов, технологий, методик исследования, применяемых в археологии к определенному типу археологических объектов;
- иерархию *Научных результатов*, позволяющую описать результаты научной деятельности археологии, такие как открытия, гипотезы, новые законы, теории, исторические факты.

Зафиксированных в онтологии научных знаний метапонятий оказалось недостаточно для отражения хронологических и географических характеристик археологических знаний. Поэтому онтология научного знания была дополнена новыми классами понятий, характерными для исторических наук:

- класс *Археологический период* служит для датирования объектов исследования и определения хронологического положения разделов науки;
- класс *Местоположение* позволяет задавать географическую привязку разделов науки (например, для разделов *Океаническая археология*, *Археология Урала*) и историко-хронологические

связи между периодами времени и их географическим местоположением (например, при задании понятий Древняя Греция или Мезолит в Северной Европе).

Контент портала представляет знания об основных разделах археологии, ее объектах, методах и результатах исследования, которые объединены по географическому и хронологическому принципу, снабжены текстовыми описаниями и ссылками на публикации. Заданная в контенте портала информация позволяет организовать содержательный доступ к ресурсам, представляющим базы данных и архивы визуальных (графических) и библиографических документов, материалов деятельности археологических организаций, отчетов экспедиций, а также описания музейных коллекций.

Портал знаний по компьютерной лингвистике [Боровикова и др., 2008] разработан для организации эффективного доступа к лингвистическим ресурсам. Пользователями такого портала могли бы стать как научные работники, преподаватели и студенты, исследующие, преподающие и изучающие эту дисциплину, так и специалисты, разрабатывающие программные системы, предназначенные для обработки текстов, анализа и синтеза речи.

В качестве онтологии области знаний в данном портале выступает онтология компьютерной лингвистики (КЛ), построенная на основе онтологии научного знания.

Для отражения в исследуемом материале компьютерной лингвистики того аспекта, на который направлена научная деятельность, онтология научного знания была дополнена новым понятием *Предмет исследования*.

В результате исследований было выделено около 200 понятий компьютерной лингвистики, которые были организованы в следующие иерархии «общее–частное»:

- иерархию *Объектов исследования*, рассматривающую следующие объекты моделирования языка: *Речевое произведение* (РП), как объективную форму существования и использования естественного языка, и *Структурные языковые единицы* в составе РП, соответствующие различным языковым уровням – предложения, словосочетания, слова, морфемы, звуки и интонационные единства.;
- иерархию *Методов исследования*, служащую для систематизированного описания инструментов исследования, применяемых в компьютерной лингвистике.;
- иерархию *Разделов науки*, в основе которой лежит классификация базовых теоретических и прикладных направлений компьютерной лингвистики;
- иерархию *Научных результатов*, служащую для типизации и описания результатов научной

деятельности. В этой иерархии были выделены такие классы: *Технологии и программные продукты, Прикладные системы, Лингвистические ресурсы*. Последний класс делится на такие классы, как *Корпуса, Лингвистические БД, Онтологии, Словари и тезаурусы*;

- иерархию *Предметов исследования*, которая описывает *Процессы*, связанные с функционированием языковых единиц в коммуникации, и *Прикладные процессы*, имеющие практическую ценность, отвечающие определенному социальному запросу.

Так как портал посвящен компьютерной лингвистике, в его контенте в первую очередь представлены знания об основных разделах компьютерной лингвистики, о ее предметах и объектах исследования, используемых в ней моделях и методах.

Ресурсы компьютерной лингвистики представлены непосредственно результатами деятельности организаций и отдельных исследователей, полученными в ходе выполнения научных и коммерческих проектов. К таким ресурсам относятся как технологии, программные продукты, прикладные системы, так и лингвистические ресурсы: словари, корпуса (текстов и речи) и лингвистические БД. Для организации более эффективного доступа к таким ресурсам в контенте представлена информация о различных аспектах их разработки: организациях, персонах и проектах, с которыми связано их появление, а также о таких содержательных характеристиках ресурсов, как отнесенность к разделу науки, объекту или предмету исследования, методам исследования. Эта информация связывает ресурсы с остальными данными и знаниями, представленными в контенте портала, что позволяет пользователю выделить группы ресурсов, созданные, например, в ходе осуществления некоторой исследовательской деятельности (гранта, проекта, конкурса) или с использованием определенного метода исследования.

4.2. Использование онтологий в системе информационной поддержки разработчиков интеллектуальных СППР

Значительно повысить скорость и эффективность создания СППР, а также их качество, может хорошая информационная поддержка разработчиков [Загорулько Г., 2013]. Она должна обеспечить разработчиков концептуальным и методологическим базисом, предоставить им систематизированные знания о научной дисциплине "Теория принятия решений" и структурированные описания методов принятия решений и решателей или пакетов, реализующих эти методы. Полученная информация позволит разработчикам лучше понять стоящие перед создаваемой СППР задачи и подобрать наиболее подходящие методы и средства для их решения.

Для того чтобы разработчик СППР смог получить удобный доступ ко всей необходимой ему

при создании СППР информации и эффективно воспользоваться ею, он, прежде всего, должен иметь представление о принципах организации и систематизации такой информации, т.е. ее концептуальной основе. В качестве концептуальной основы информационной поддержки разработчика предлагается использовать онтологию задач и методов принятия решений (ЗиМППР). Рассмотрим, какими особенностями должна обладать онтология ЗиМППР.

Порталы научных знаний, рассматриваемые в предыдущем разделе, ориентированы на обслуживание одного типа пользователей, например, ученого или производственника. Этим определялись и структура онтологий портала, и требования к ним. Однако существует потребность в создании систем, отражающих интересы различных типов пользователей. Поэтому формализация семантики таких областей знаний должна учитывать несколько типов пользователей.

Одной из таких проблемных областей является разработка СППР. В этой области мы имеем, по крайней мере, три типа пользователей – эксперт, инженер знаний, программист. Каждый из этих пользователей на своем уровне выступает как ЛПР, и каждому из них нужна такая же поддержка, какая осуществляется порталом знаний. Но ситуация здесь осложняется тем, что мы как и прежде строим одну систему информационной поддержки, но рассчитанную на разные группы пользователей одновременно. Это ведет к существенному усложнению системы знаний, которая не может строиться путем включения отдельных онтологий для каждого типа пользователей, так как эти онтологии имеют общих частей не меньше, чем различных. И если эти онтологии строить независимо одна от другой, то это может привести к противоречиям в системе понятий, а в дальнейшем и к противоречивости всей системы знаний. В связи с этим в рассматриваемой в этом разделе системе информационной поддержки была предложена структура онтологии, отличная от структуры онтологий для традиционных ПНЗ.

Для того, чтобы свести профессиональную терминологию всех типов пользователей к общей системе понятий ЗиМППР должна иметь так называемую «вертикально-горизонтальную» структурную организацию, что позволит представлять интересующую область знаний в двух измерениях (см. рисунок 6).

Первое измерение («вертикальная структуризация») представляется в виде традиционной иерархической структуры, на каждом уровне которой интересующая область знаний описывается с разной степенью детализации.

Второе измерение («горизонтальная структуризация») задает описание области знаний с точки зрения разных типов специалистов, участвующих в процессе создания и использования СППР.



Рисунок 6 – Вертикально-горизонтальная организация онтологии ЗиМППР

В качестве общего ядра, на котором строится вертикально-горизонтальная онтология ЗиМППР, выступает метаонтология задач и методов, которая содержит описание таких базовых понятий поддержки принятия решений, как Задача, Метод, Модуль, Решатель, Входные данные, Результат, Ситуация, Проблемная ситуация, Альтернатива, Этап принятия решений, а также отношения между ними [Загорулько Г., 2013а].

Онтология ЗиМППР тесно связана с понятиями предметных областей, для которых создаются СППР. Для того чтобы дать представление о задачах и методах, не вдаваясь в конкретику ПО, описания этих областей должны быть одинаково устроены. Это задается путем использования в качестве основы для построения онтологии конкретной предметной области метаонтологии ПО, включающей базовые понятия, являющиеся общими для всех ПО.

Для реализации информационной поддержки разработчиков СППР были использованы средства технологии создания порталов научных знаний, которая позволяет поддержать изложенные в докладе концептуальные и методологические принципы обеспечения информационной поддержки. На основе онтологии, описывающей предметную область, она позволяет строить структурированный контент, представляющий собой семантическую сеть объектов, привязанных к понятиям онтологии ПО, где предметная область – это поддержка принятия решений, а объекты – это конкретные задачи, методы, этапы и т.д.

Заключение

В докладе показано, что формализация семантики области знаний в виде онтологий – это весьма продуктивный подход. Он позволяет эффективно представлять как объективные знания (ориентированные на машинную обработку), так и субъективные знания (ориентированные на различные группы пользователей), необходимые для функционирования информационных и интеллектуальных систем.

Было отмечено, что к настоящему времени разработано большое число средств различного

уровня для формализации семантики – логики, языки, редакторы, среды.

В докладе также показано, что онтологии могут использоваться как при проектировании и разработке информационных и интеллектуальных систем, так и в качестве полноправного компонента во время функционирования системы.

При разработке информационных и интеллектуальных систем онтология может использоваться для:

- формирования и фиксации общего разделяемого всеми экспертами знания о предметной области;
- описания экспертных правил и прецедентов;
- явной концептуализации ПО, позволяющей описывать семантику данных;
- обеспечения возможности переиспользования знаний;
- описания функциональности системы (типов решаемых задач);
- разработки отдельных компонентов системы (в качестве их высокоуровневой спецификации);
- разработки хранилищ данных (в качестве высокоуровневой спецификации структуры хранилища);

В процессе функционирования систем перечисленного класса онтология может применяться для обеспечения:

- совместного использования разнородных данных и знаний в рамках одной системы;
- процесса решения задач, составляющих функциональность системы;
- лучшего понимания предметной области пользователями системы.

Работа выполнена при финансовой поддержке РФФИ (проект № 13-07-00422) и Президиума РАН (интеграционный проект СО РАН № 15/10 «Математические и методологические аспекты интеллектуальных информационных систем»)

Библиографический список

[Андреева и др., 2006] О.А. Андреева, О.И. Боровикова, С.В. Булгаков, Ю.А. Загорулько, Е.А. Сидорова, Б.Г. Циркин, Ю.П. Холмошкин. Археологический портал знаний: содержательный доступ к знаниям и информационным ресурсам по археологии // Труды 10-й национальной конференции по искусственному интеллекту с международным участием - КИИ'2006. – Москва: Физматлит, 2006. –Т.3. С.832-840.

[Ануреев и др., 2009] Ануреев И.С., Батура Т.В., Боровикова О.И., Загорулько Ю.А., Кононенко И.С., Марчук А.Г., Марчук П.А., Мурзин Ф.А., Сидорова Е.А., Шилов Н.В. Модели и методы построения информационных систем, основанных на формальных, логических и лингвистических подходах / Отв. ред. А.Г. Марчук ; Рос. акад. наук, Сиб. отд-ние, Ин-т систем информатики им. А.П. Ершова. – Новосибирск: Изд-во СО РАН, 2009. ISBN 978-5-7692-1113-3. – 330 с.

[Боровикова и др., 2008] Боровикова О.И., Загорулько Ю.А., Загорулько Г.Б., Кононенко И.С., Соколова Е.Г. Разработка портала знаний по компьютерной лингвистике // Труды 11-ой национальной конференции по искусственному

интеллекту с международным участием КИИ-2008. М.: ЛЕНАНД, 2008. Т.3. С.380-388.

[**Варшавский и др., 2006**] Варшавский, П.Р. Методы правдоподобных рассуждений на основе аналогий и прецедентов для интеллектуальных систем поддержки принятия решений / П.Р. Варшавский, А.П. Еремеев // *Новости Искусственного Интеллекта*, № 3, 2006, С. 39-62.

[**Загорулько Ю.А. и др., 1996**] Попов И.Г.. Представление знаний в интегрированной технологической среде Semp-ТАО. // *Проблемы представления и обработки не полностью определенных знаний*. –Москва-Новосибирск, 1996. –С.59–74.

[**Загорулько и др., 2008**] Загорулько Ю.А., Боровикова О.И. Подход к построению порталов научных знаний // *Автометрия*. № 1, 2008, т. 44. С. 100–110.

[**Загорулько и др., 2010**] Загорулько Ю.А., Загорулько Г.Б., Кравченко А.Ю., Сидорова Е.А. Разработка системы поддержки принятия решений для нефтегазодобывающего предприятия // *Труды 12-й национальной конференции по искусственному интеллекту с международным участием – КИИ-2010*. – Москва: Физматлит, 2010. Т.3. С.137-145.

[**Загорулько, 2013**] Загорулько Ю.А. О концепции интегрированной модели представления знаний // *Известия Томского политехнического университета*. – 2013. – Т. 322. – № 5. С. 98-103.

[**Загорулько, 2013а**] Загорулько Ю.А. Технологии разработки интеллектуальных систем, основанные на интегрированной модели представления знаний // *Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2013): материалы III Междунар. научн.-техн. конф. (Минск, 21-23 февраля 2013 г.) / редкол.: В. В. Голенков (отв. ред.)*. – Минск: БГУИР, 2013. С. 31–42.

[**Загорулько и др., 2013б**] Ю.А. Загорулько, Г. Б. Загорулько, В.К. Шестаков, И.С. Кононенко. Концепция и архитектура тематического интеллектуального научного интернет-ресурса // *Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды XV Всероссийской научной конференции RCDL'2013, Ярославль, Россия, 14-17 октября 2013 г. – Ярославль: ЯрГУ, 2013. С.57–62.*

[**Лапшин, 2010**] Лапшин В.А. Онтологии в компьютерных системах. М.: Научный мир, 2010. – 224 с.

[**Загорулько Г., 2013**] Загорулько Г.Б. Обеспечение информационной поддержки разработчиков СППР // *Информационные и математические технологии в науке и управлении / Труды XVIII Байкальской Всероссийской конференции «Информационные и математические технологии в науке и управлении»*. Часть III. – Иркутск: ИСЭМ СО РАН, 2013. С. 137–142.

[**Загорулько Г., 2013а**] Загорулько Г.Б., Загорулько Ю.А. Подход к интеграции разнородных методов поддержки принятия решений для сложных задач // *Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2013): материалы III Междунар. научн.-техн. конф. (Минск, 21-23 февраля 2013 г.) / редкол.: В. В. Голенков (отв. ред.)*. – Минск: БГУИР, 2013. С. 265–268.

[**Лукашевич, 2011**] Лукашевич Н.В. Тезаурусы в задачах информационного поиска. –М.: Изд-во МГУ, 2011.

[**Мартынов, 1974**] Мартынов В.В. Семитологические основы информатики. –Минск: Наука и техника, 1974. –192 с.

[**Нариньяни и др., 1998**] Нариньяни, А.С. Программирование в ограничениях и недоопределённые модели / А.С. Нариньяни, В.В. Телерман, Д.М. Ушаков, И.Е. Швецов // *Информационные технологии*, 1998. – №7. С. 13-22.

[**Петровский, 2009**] Петровский А.Б. Теория принятия решений. Москва, Издательский центр «Академия», 2009.

[**Попов, 1987**] Попов Э.В. Экспертные системы: Решение неформализованных задач в диалоге с ЭВМ. –М.: Наука. 1987. – 288 с.

[**Тузовский и др., 2005**] Тузовский А. Ф., Чириков С. В., Ямпольский В. З. Системы управления знаниями (методы и технологии). – Томск: Изд-во НТЛ, 2005. 260 с.

[**Vaader et al., 2003**] Vaader, F. The Description Logic Handbook: Theory, Implementation, Applications / F. Vaader, D. Calvanese, D.L. McGuinness, D. Nardi, P.F. Patel-Schneider – Cambridge, 2003. 574 P.

[**Benjamins et al., 1998**] Benjamins V.R. and Fensel D. Community is Knowledge! in (KA)2 // *Proceedings of the 11th Banff*

Knowledge Acquisition for Knowledge-based Systems workshop, KAW'98 (Banff, Canada, April 1998). – Calgary: SRDG Publications, Department of Computer Science, University of Calgary, 1998.

[**Cabral, 2004**] L Cabral, J Domingue, E Motta, T Payne, F Hakimpour. Approaches to semantic web services: an overview and comparisons // *The semantic web: Research and applications*, 2004. p. 225-239.

[**Erling, 2009**] Erling O., Mikhailov I. RDF Support in the Virtuoso DBMS // *Networked Knowledge-Networked Media*. – Springer Berlin Heidelberg, 2009. p. 7-24.

[**Gruber, 1995**] Gruber T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // *International Journal of Human-Computer Studies*. November 1995. Vol. 43. Issues 5–6. P. 907–928.

[**Guarino, 1998**] Guarino N. Formal Ontology in Information Systems // *Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, June 6–8, 1998 / Ed. N.Guarino*. Amsterdam: IOS Press, 1998. P. 3–15.

[**Hillmann., 1995-2013**] Hillmann D. Using Dublin Core. <http://dublincore.org/documents/usageguide/> (дата обращения: 14.12.2013)

[**OWL-S, 2004**] OWL-S: Semantic Markup for Web Services. <http://www.w3.org/Submission/OWL-S/> (дата обращения: 14.12.2013)

[**OWL, 2013**] OWL Web Ontology Language Guide. W3C Recommendation 10 February 2004. URL: <http://www.w3.org/TR/owl-guide/> (дата обращения: 14.12.2013).

[**Protégé, 2013**] Protégé. URL: <http://protege.stanford.edu/> (дата обращения: 14.12.2013).

[**SWRL, 2013**] SWRL: A Semantic Web Rule Language Combining OWL and RuleML. W3C Member Submission 21 May 2004. <http://www.w3.org/Submission/SWRL/> (дата обращения: 14.12.2013).

[**Tarski, 1956**] Tarski A. Logic, Semantics, Metamathematics. – Oxford University Press, 1956. – 258 p.

[**Virtuoso, 2013**] Virtuoso open-source edition. OpenLink Software. <http://www.openlinksw.com/wiki/main> (дата обращения: 14.12.2013)

ONTOLOGY-BASED FORMALIZATION OF SEMANTICS OF KNOWLEDGE AREAS IN INFORMATION AND INTELLIGENT SYSTEMS

Zagorulko Yu.A. *, Zagorulko G.B. *

** A.P. Ershov Institute of Informatics Systems
Siberian Branch of the Russian Academy of
Sciences, Novosibirsk, Russia*

zagor@iis.nsk.su

gal@iis.nsk.su

The paper discusses the role of ontology in intelligent and information systems, the ontology-based approaches to and means for formalizing the semantics of knowledge areas are presented and the examples of practical use of ontologies in systems of this class are discussed.

The paper shows that the formalization of the semantics of the knowledge area in the form of ontologies is a very productive approach It allows you to effectively represent both objective knowledge (oriented to machine use) and subjective knowledge (targeted at different groups of users) of the modeled area, necessary for the operation and information intelligent systems.

In conclusion, it is concluded that the ontology can be used successfully at both development time and run time of information and intelligent system.