



OSTIS-2014

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822

СЕМАНТИКА И ОНТОЛОГИИ В ЕСТЕСТВЕННЫХ И ИСКУССТВЕННЫХ СИСТЕМАХ

Сторож В.В.

*Донецкий государственный технический университет,
г. Донецк, Украина
ws52@mail.ru*

В работе рассмотрено влияние моделей семантики и созданных на их базе онтологий на возможности современных интеллектуальных систем. Показано, что широко распространенные в настоящее время модели семантики имеют критически важные ограничения по поиску, интеграции и анализу информации. Предложено использовать интеллектуальные системы с моделями семантики, в большей степени учитывающими специфику анализа информации человеком. Сформулирован ряд требований к составу и содержанию таких систем.

Ключевые слова: требования к моделям семантики; сущность онтологий; учет принципов человеческого мышления.

Введение

Большие надежды в области поиска, интеграции и анализа информации в последние годы связаны с онтологиями, как одним из направлений искусственного интеллекта (ИИ). При этом особенности реализации и возможности онтологий существенно зависят от модели семантики, принятой в данной онтологии [Colomb, 2007]. В то же время используемые в ИИ модели семантики существенно беднее семантики человека [Поспелов, 1989; Кузнецов, 2012; Васильев, 1990; Попова и др., 2007]. Очевидно, что неадекватность моделей семантики семантике человека должна привести к каким-то ограничениям на использование онтологий. Каковы же эти ограничения? По-видимому, этот вопрос никем систематически не исследовался.

Существует не только теоретическая, но и настоящая практическая потребность в понимании того, какие ограничения накладывают используемые семантические модели на онтологии, и как можно снять, либо хотя бы ослабить эти ограничения. В качестве примера приведем Интернет-технологии Semantic Web, основанную на онтологиях, и широко разрекламированную Тимом Тернерсом-Ли в 2001 году [Terners-Lee et al., 2001]. Интернет сообщество и ведущие разработчики программного обеспечения с энтузиазмом подхватили эту технологию и начали внедрять ее в работу Интернет и в программные продукты. Но достижения пока довольно скромные, что признал и сам автор Semantic Web [Shadbolt et al., 2006].

Радикальные идеи, как исправить ситуацию, в настоящее время отсутствуют.

Было бы наивно полагать, что поднятые выше вопросы можно решить и осветить в одной публикации. Целью данной работы является скорее не развернутый ответ на поставленные вопросы, а междисциплинарное рассмотрение использования семантики в онтологиях, и на этой основе более четкая формулировка исходной ситуации, направлений исследований и области возможных применений этих исследований.

1. О семантике в различных дисциплинах

Слово «семантика» является понятием (синоним – концепт), и по общему свойству всех понятий обладает объемом и содержанием [Попова и др., 2007]. Рассмотрим ключевые составляющие этого объема и содержания в системах ИИ и у человека (такой порядок рассмотрения выбран для большей наглядности и уменьшения повторов).

Общие принципы формальной семантики в системах ИИ приведены в [Голенков и др., 2012]. Они сводятся к следующему:

«В качестве формальной основы проектируемых интеллектуальных систем, в качестве основы абстрактных логико-семантических моделей интеллектуальных систем, использовать графодинамические модели специального вида – семантические модели представления и обработки знаний, в основе которых лежат семантические сети.

...все элементы (атомарные фрагменты) семантической сети являются знаками различных сущностей (объектов). Такими сущностями могут быть всевозможные внешние описываемые объекты, а также различные множества, состоящие из элементов (атомарных фрагментов) этой же семантической сети.

...Таким образом, семантическая сеть – это абстрактная знаковая конструкция «рафинированного вида», в которой нет ничего кроме знаков и инцидентности этих знаков».

Имеется целый ряд более конкретных трактовок формальных семантических моделей. В частности [Colomb, 2007, p.29] выделяет:

- отношения сущность-связь (Entity-Relationship model), где используются понятия сущности, экземпляра, отношения, атрибута и подтипа;
- унифицированный язык моделирования (Unified Modeling Language – UML), использует терминологию класса, экземпляра, ассоциации, атрибута и подкласса;
- объектно-ролевое моделирование (Object-Role Modeling), использует терминологию типа объекта, объекта, типа факта, роли и подтипа;
- дескриптивная логика (Description logics), использует терминологию концепта, индивидуального, роли и включения;
- язык веб онтологий (Web ontology Language – OWL), использует терминологию класса, индивидуального, свойства и подкласса;
- общая логика (Common Logic – CL), основана на синтаксисе исчисления предикатов первого порядка и использует терминологию выражение, атомарный факт, предикат и функция.

Несмотря на существенные различия перечисленных семантических моделей, в них можно выделить и общие черты:

- понятия, а если точнее, модели понятий, являются одним из ключевых элементов во всех типах семантических моделей. В некоторых типах моделей, в частности, логических, модели понятий являются единственным типом ключевых элементов;
- модели понятий могут содержать в себе более мелкие информационные единицы: атрибут, свойство, значение и др. Причем максимальное количество уровней детализации не превышает двух (сущность-атрибут-значение);
- возможны различные типы взаимоотношений между различными моделями понятий, а также их компонентами, в частности: включение в класс, роль, тип объекта, следование и др.;
- включение ключевого элемента в некоторую общность более высокого уровня не приводит к автоматическому изменению характеристик этого ключевого элемента (это утверждение не включает представление ключевых элементов с помощью объектно-ориентированного программирования, а также как агентов). Этот тезис можно

сформулировать и по принципу «сверху-вниз»: возможна декомпозиция некоторого информационного объекта, более крупного, чем модель понятия, на более мелкие элементы без изменения информационного содержания как системы в целом, так и составляющих ее элементов. Еще одна трактовка – информационные объекты являются линейно независимыми.

Рассмотрим понятие семантики в лингвистике. Сразу отметим, что в лингвистике мы также имеем дело с моделями. Однако эти модели получены в результате лингвистического анализа носителей семантики (людей) и поэтому они являются более адекватными, чем модели семантики в ИИ. К недостаткам этих моделей относится низкий уровень формализации, в результате чего они только описывают ситуацию, но не пригодны без дальнейшей формализации к использованию в автоматизированных системах анализа информации (если только не «вложить» эти модели в носителя семантики – человека).

С точки зрения лингвистов [Васильев, 1990, с.67] «Многочисленные определения значения, предлагавшиеся представителями различных направлений языкознания, можно свести к двум основным:

- 1) К определению значения как психической, отражательной сущности, соотносимой с такими психическими явлениями, как представления, эмоции, понятия и т.п.
- 2) К определению значения как реляционной сущности, как отношения языкового знака к предмету (денотативное значение), к понятию (сигнификативное значение), к условиям акта речи и его участникам (прагматическое значение), к сфере или ситуации его употребления (стилистическое значение), к другим языковым знакам (структурное значение) и т.д.».

Очевидно, что все без исключения модели семантики в ИИ не содержат в себе психических переживаний типа эмоций, представлений, сознания и т.п. – т.е., всего, что описывается лингвистами в п.1.

В то же время некоторые типы значений (семантики), приведенные в п.2 содержатся, или могут содержаться при определенных условиях, и в моделях семантики, используемых в ИИ. В частности, структурное значение семантики в ИИ и у человека практически совпадают. Для реализации прагматического значения в интеллектуальной системе необходимо реализовать модель мира или окружающей среды, а также цели и задачи субъекта порождения семантики. Такого рода условия моделируются, в частности, в экспертных системах и некоторых СППР. Поэтому можно сказать, что прагматическая семантика может частично содержаться в системах ИИ.

Для отображения стилистических значений необходимо моделирование текущей ситуации,

субъектов влияния, их действий, намерений и т.д. По-видимому, такая задача в настоящее время ИИ недоступна (если рассматривать уровень людей, а не примитивных искусственных агентов), хотя принципиальных препятствий для ее реализации нет, и она может быть решена в обозримом будущем.

В [Попова и др., 2007, С.94-100] различается три вида значения понятия или концепта:

1) Лексикографическое – это то, что фиксируется в словарях. Именно это значение используют специалисты по ИИ, фиксируя его в базах знаний и в онтологиях.

2) Психолингвистическое – это весь набор значений, который могут зафиксировать лингвисты у носителей языка в результате углубленных исследований. Психолингвистическое значение гораздо шире и объемней, нежели его лексикографический вариант, который составляет лишь небольшую часть психологически реального значения.

3) Но содержание понятия (концепта) еще шире психолингвистического значения, поскольку оно включает в себя смысл как неформализуемую, т.е. не выражаемую словесно часть восприятия.

Выше были рассмотрены концепции семантики естественного языка, в которых ключевые единицы семантики либо не выделяются [Васильев, 1990], либо в качестве таких единиц используются понятия (концепты) [Попова и др., 2007]. Но объектами семантики являются не только концепты! Так, в Большом энциклопедическом словаре [Ярцева, 1989. – С.438-439] семантика представляется в виде многоуровневой иерархической системы семантических значений, включая:

- уровень морфем (семантика морфемы);
- уровень понятий или концептов (семантика концепта);
- уровень предложений (семантика предложения);
- уровень текста (семантика текста).

И что самое существенное, семантика, т.е. содержание каждого уровня не является просто суммой (или частью) нижестоящих (вышестоящих) уровней [Васильев, 1990, с.70-71.].

Нагляднее всего это положение можно проиллюстрировать на примерах. Так, выражение «сидеть на солнце» (идея примера взята из [Кобозева, 2000]), вовсе не означает сидеть, свесив ноги, на центральном объекте Солнечной системы. Оно предполагает субъекта или субъектов, солнечную погоду, отдых и т.д., т.е., довольно развитую модель ситуации, которая в ряде своих составляющих никаким образом не содержится в определении исходных понятий.

Второй пример: пусть некто в шутку или всерьез заявляет, что у них в организации монголо-татарское иго. Если следовать буквальному смыслу

входящих в выражение понятий, то в организации имеются принадлежащие монголам и татарам хомуты. При выходе на первый вышестоящий уровень мы обращаемся к истории, и понимаем это как захват Руси монголо-татарами. На втором уровне мы понимаем, что угнетение, по тяжести сравнимое с монголо-татарским игом, происходит в данной организации. Таким образом, входящие в состав выражения понятия переосмысливаются коренным образом дважды, и вывод из данного выражения делается не на основании значения входящих в него понятий, а на основании вновь сконструированного смысла.

Из проведенного сравнительного обзора формальной и естественной семантики можно сделать следующие выводы:

- формальная семантика использует лишь небольшую часть значения понятий по сравнению с носителями естественного языка;
- в формальной семантике значения вышестоящих уровней формируются из нижестоящих по неким типовым законам. При этом содержание единиц нижестоящего уровня остается практически неизменным;
- в семантике естественных языков содержание вышестоящих уровней формируется каждый раз индивидуальным образом с учетом истории, традиций, культуры вообще, а также с учетом психологических особенностей субъекта восприятия, - его намерений, эмоционального состояния и др. При этом значение единиц нижестоящего уровня может существенно изменяться;
- в формальных системах для получения выводов и принятия решений фактически используется информация одного-единственного уровня. Получение вывода и семантика есть почти одно и то же;
- человек вначале создает на основании речи или языка содержание верхних уровней семантики, а затем использует его для получения выводов и принятия решений. Конструирование семантики и получение на ее основе выводов, - это разные (но последовательные) психические процессы.

В таблице 1 суммированы сравнительные характеристики объема семантики в естественном и искусственном интеллектах.

Таблица 1 – Сравнительные характеристики объема семантики

Компонента семантики	ИИ	Человек
Прагматика	+/-	+
Стилистика	+--	+
Структурное значение	+	+
Психологическое	-	+
Лексикографическое	+	+
Психолингвистическое	+--	+
Смысл	-	+
Нелинейность	-	+
Иерархия уровней	-	+

Обозначения: +- - качество частично присутствует в некоторых современных разработках;

+-- - качество не реализовано в современных разработках вследствие высокой трудоемкости, но принципиальных ограничений для его реализации нет;

- - качество принципиально невозможно реализовать в системах ИИ на основании используемых принципов.

Сравнение объема семантики в ИИ и естественном интеллекте показывает следующее:

- из целого ряда семантических компонент ИИ полноценно использует лишь одну – структурную, которую с равным успехом можно назвать не семантической, а синтаксической;

- в ИИ в принципе не присутствуют и не создаются значения вышестоящих уровней, которые, собственно и используются для получения выводов.

В качестве комментария: такое тяжелое психическое заболевание как аутизм с когнитивной точки зрения заключается в буквальном понимании слов [Питерс, 1992, с.29-30], т.е. в неспособности строить семантические значения как для понятий, так и для высших уровней. И такого посредника мы создаем между собой и необозримыми базами знаний. На основании проведенного рассмотрения оптимизм по поводу творческих способностей такого «интеллекта» представляется не просто необоснованным, а очевидно неверным.

И в заключение данного вопроса о нейрофизиологических «корнях» семантики и сущности формализации.

Мозг является самой сложной системой во Вселенной, известной человеку. Нейрофизиологи выделяют не менее девяти уровней его структурной организации [Shepherd, 2004, р.6]. Поскольку каждый структурный уровень задает какие-то свои специфические информационные процессы, то число уровней представления информации (и соответствующих этим уровням информационных сущностей) в мозгу также не менее девяти.

На рис.1. представлены примерные соответствия между уровнями организации мозга и компонентами рассмотренных выше типов семантики.

Самые дробные информационные единицы, с которыми работают формальные методы, - это признаки или значения. С точки зрения нейрофизиологии этому соответствует уровень никак не меньше нейрона (а скорее всего даже на один-два уровня выше). Поскольку же формальные методы не способны к построению семантики на вышестоящих уровнях информационного содержания, то они неспособны и к генерации целостных информационных объектов вышестоящих уровней. Т.е., их уровень

использования целостности близок нулю. Эта ситуация и представлена на рис.1.

С другой стороны, смысловое информационное содержание интегрирует в некоторую целостность всю (или значительную часть) доступной человеку информации, используя для этого все структурные и процессные уровни представления информации (что также представлено на рис.1.).

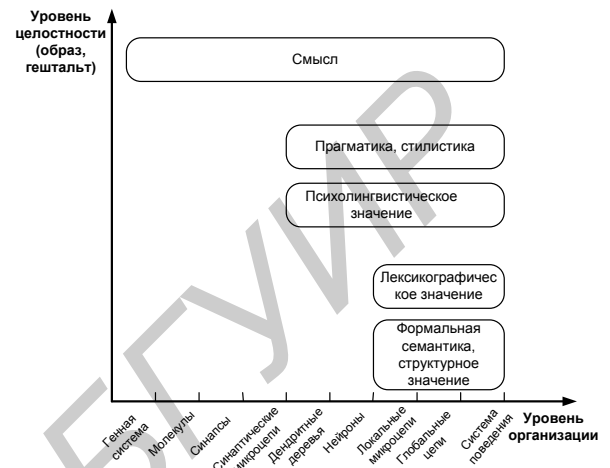


Рис.1. Взаимосвязь типов семантики с уровнями организации мозга

Опираясь на предыдущее рассмотрение, и наглядное представление на рис.1, можно сделать чрезвычайно важный для последующих рассуждений вывод: сущность формализации заключается в отбрасывании ряда нижележащих уровней обработки информации. Результаты их, вообще говоря, сложной индивидуальной деятельности, заменяются существенно более простыми и однотипными отношениями (логическими выражениями или математическими функциями). Как показывают оценки [Сторож, 2013], такая процедура позволяет уменьшить информационные потоки в ИИ на 20 и более порядков. Но неизбежно возникает вопрос: а что же при этом теряется? Несмотря на исключительную важность ответа на этот вопрос, специалисты по ИИ его, по-видимому, не исследовали. Проведенные же в данной работе исследования объема потерь «по факту» показывают, что теряется смысл, целостность и компоненты семантики, участвующие в понимании и мышлении.

2. Онтологии и формальные семантики

Что такое онтология (или онтологии) в ИИ? Дадим несколько определений, в которых формулируются различные взгляды на сущность этого понятия. Наиболее известно определение, восходящее к Груберу, и затем уточненное Студером и др. [Studer, 1998]: «Онтология является формальной, явно выраженной спецификацией совместной концептуализации».

Широко известно и цитируется определение Нариньяни А.С. [Нариньяни, 2001]: «Онтология, это модель предметной области, использующая все доступные средства представления знаний, релевантные для данной области».

В связи с развитием Semantic Web возникло еще одно уточняющее определение [Ландэ, 2005]: «Онтология – это «спецификация концептуализации предметной области», или упрощенно документ или файл, формально задающий отношения между терминами. Это своего рода словарь понятий предметной области и совокупность явным образом выраженных предположений относительно смысла этих понятий».

Прокомментируем эти определения. Первое из них выражает позицию разработчиков программного обеспечения и содержит целый ряд спорных, узких или недостаточно определенных положений.

Так, в [Guarino, 2009, p.2] отмечается, что термин «концепт» в данном случае является существенно неадекватным, поскольку он не несет в себе смысла сущности, существования чего-либо, и предлагают заменить его более уместным, в частности, термином «универсалия».

Вызывает вопросы и использование словосочетания «спецификация совместно разделяемых знаний», поскольку под спецификацией принято обычно понимать некоторый перечень, список. Но к формальным совместно разделяемым знаниям могут относиться также и программы, схемы, графические и процессные модели. Более того, эти типы представления знаний широко используются в настоящее время для создания онтологий. Поэтому термин «спецификация» необходимо заменить на термин с более широким значением, включающий все типы формального представления знаний.

И третье замечание к определению Грубера-Студера рассмотрим в контексте определения, данного А.С. Нариньяни. Онтология является не просто совместным представлением чего-либо, а таким представлением, в котором выделяются наиболее существенные свойства предметной области, - по определению, это модель. Но эту мысль можно выразить еще более определенно: создавая модель, мы выделяем инварианты, описывающие предметную область и позволяющие наиболее экономным и оптимальным образом описывать ее с точки зрения задач, которые мы ставим по отношению к данной предметной области. В частности, эта идеология и терминология широко используется в таком направлении кибернетики, как идентификация систем.

И относительно определения, связанного с Semantic Web. В нем подчеркивается, что онтология задает не только список понятий, связанных с какой-либо предметной областью, но и отношения между понятиями, смысл. Иначе говоря,

неотъемлемой, центральной частью любой онтологии является та или иная модель семантики, принятая данной онтологией. И именно принятая в онтологии модель семантики определяет тип языков представления знаний, используемых моделей, и, в конечном счете, область использования и возможности данной онтологии.

С учетом приведенных определений и обсуждения их содержания, дадим свое определение онтологии:

Онтология, - это инвариантное, наиболее экономное и последовательное с точки зрения выполняемых задач представление содержания некоторой предметной области или вида деятельности, основанное на использовании определенной модели семантики, и одного или нескольких стандартизированных языков представления знаний.

В это определение не включено требование разделяемости онтологии участниками, поскольку на этапе разработки онтологии оно просто невозможно, а в случае успеха онтологии обеспечивается автоматически.

Тезис об использовании в онтологии стандартизированных языков представления знаний включен в определение на основании следующих соображений:

- во-первых, все наиболее известные онтологии основаны на ранее созданных или разработанных для данной онтологии языках представления знаний [Colomb, 2007, p.29];
- во-вторых, использование в онтологии стандартизированного языка представления знаний существенно облегчает задачу изучения данной онтологии сообществом пользователей, т.е. ее «совместную разделяемость».

В нашем определении онтология основывается на содержании не только предметной области, но и вида деятельности. Это связано с тем, что понятие «предметная область» связывается в основном с текущим положением, статикой. А реальные онтологии во многих случаях в той или иной мере предназначены для описания действий, т.е. динамики. В частности, это относится к Semantic Web, идеологии онтологического проектирования, и вообще всем типам моделей с участием процессов, работ, действий и т.п.

3. Требования к моделям семантики и онтологиям

Переформулируем приведенное выше определение онтологий с учетом их практических функций:

Входящие в онтологии информационные объекты представляют собой своеобразные интеллектуальные точки отсчета, при использовании которых некоторый класс задач, входящих в сферу действия данной онтологии,

решается максимально быстрым и экономным образом.

Каким образом это происходит? Для иллюстрации приведем несколько широко известных интеллектуальных объектов, выполняющих те же функции, что и онтологии.

В механике для упрощения решения уравнений движения используются преобразования координат фазового пространства. Фактически ищется некоторая новая система координат, в которой уравнения решаются максимально просто [Ланцош, 1965, с.227].

В многомерной статистике существует сходный принцип в методе выделения главных компонент, приводящий к уменьшению числа переменных. Он заключается в последовательном вращении переменных с целью максимизации дисперсии исходного пространства переменных, а полученные направления используются как главные компоненты [Айвазян, 1998, с.565].

Т.е., ключевой принцип, заложенный в онтологиях, не является для человечества совершенно новым, он широко использовался и ранее. Но в отличие от рассмотренных примеров, онтологии работают с областями знаний, для которых процедуры формальной оптимизации заключаются не во вращении координат, а в выборе ключевых точек в концептуальном пространстве, и связей между ними. Причем такая процедура, по-видимому, не является полностью формализуемой, хотя и включает использование математики, например, для определения частоты связей между понятиями.

Полученные с использованием вышеописанных принципов онтологии могут использоваться в двух ключевых режимах:

1) В режиме выполнения стандартных операций или процедур. Очевидно, что для организации такой деятельности достаточно формальной семантики, заложенной в онтологиях (конечно, если она адекватна выполняемым задачам).

2) В режиме, где недостаточно стандартных знаний, и требуются новые идеи, инновации, разработки и т.п. Онтологии также выполняют полезную роль, позволяя сформулировать оптимальную для поиска исходную интеллектуальную позицию. Но самих по себе онтологий для получения правильного решения, как правило, недостаточно, что подтверждается на опыте.

Почему это происходит? С самых общих позиций можно сказать, что в онтологиях не хватает гибкости, «разрешающей способности». Новую идею, видение, понимание, можно получить, используя неожиданные, ранее не формализованные связи между понятиями и группами понятий, которых в онтологиях нет по определению, нет в них и генерации новых значений на уровне групп

понятий (выражений, предложений), и более крупных участков текста. А именно на этом уровне происходит понимание конкретных ситуаций, рождение новых идей, и вообще всех интеллектуальных процессов, недоступных компьютеру.

Кто-то может сказать, что онтологии и не ориентированы на автоматический вывод новых идей, но вот для поиска информации они подходят превосходно. Но любой специалист прекрасно знает, что это не так. На большинство вопросов поисковики выдают огромное количество ссылок (обычно от десятков тысяч до миллионов). Выделить среди них нужную информацию средствами Интернет невозможно, вот и приходится ограничиваться первыми десятками или сотнями, пропуская, может быть критически важную информацию. Почему формальные методы не могут отобрать нужную информацию? Можно выделить целый ряд причин:

- новые идеи используют и новые связи между понятиями. Естественно, этих новых связей статистически немного по сравнению с традиционными, и они просто не попадают в онтологии. Их появление может заметить лишь человек в результате углубленного изучения информации;

- одни и те же понятия различными людьми понимаются существенно различным образом; и, наоборот, называемые разными словами представления могут быть близки между собой. Поэтому человек при углубленном анализе информации «индивидуализирует» понимание ключевых понятий в тексте в зависимости от их происхождения;

- при поиске информации в Интернет компьютерными методами неявно предполагается, что вся информация правда. Нет нужды доказывать, что это далеко не так. Информация может быть искажена самыми различными способами: прямая ложь, преувеличение в рекламных целях, умолчание в той или иной степени, подмена понятий; пересказы известного, в той или иной мере неточные (просто информационный мусор) и др.

Не подходят современные формальные методы и для аналитики уже готовых отобранных текстов (например, базы знаний предприятия). Как уже говорилось, многие (если не подавляющее большинство) ключевые идеи выражаются не перешедшими в штампы словосочетаниями, а на этом уровне формальные методы семантику просто не улавливают.

На основании вышеизложенного можно сделать вывод, что повышение эффективности интеллектуальных систем на основе онтологий напрямую зависит от уровня семантики этих интеллектуальных систем. Причем уровень семантики определяется не мощностью формальных методов, а соответствием человеческому мышлению, тем принципам, которые использует человек при генерации новых идей,

оценке ситуаций, построении моделей конкретных объектов и действий и т.д. Человек не просто использует стандартные модели семантики, он в каждом конкретном случае «строит» семантику, создавая смыслы верхних уровней. И эти уровни сами по себе никаким образом в тексте не содержатся. Для их генерации необходим мощный интеллект, владеющий культурой, и способный понимать ситуации, к которым относится текст или любая другая не полностью формализованная информация.

Под интеллектом в данном случае понимается программная конструкция типа экспертной системы, но с существенно более широкими возможностями. Эта конструкция должна включать в себя систему порождения смысла высших уровней, и получения выводов с использованием этих смыслов. Для этого необходимы:

- база знаний самого интеллекта (в отличие от анализируемой информации), включающая хотя бы в некоторой степени культуру, историю, и интерпретацию (смысл, семантику) стандартных словосочетаний;
- способность к порождению смыслов высших уровней, используя, в частности, механизмы построения метафор и аналогий;
- механизмы интеграции информации не на основе слов, а на основе порождаемых смыслов;
- механизмы оценки качества информационного содержания и на этой основе присваивание рейтинга тексту;
- человек не использует при мышлении сложных функций, а обработка информации происходит существенно параллельным образом, и всего за несколько циклов. В связи с этим необходимо искать и использовать простые и в высокой степени параллельные механизмы порождения семантики.

4. Перспективные направления использования онтологий

Интеллектуальные системы с использованием онтологий можно улучшать двумя принципиально различными способами:

1) Разработка некоторого формального языка, в семантике которого максимально возможным образом заложены и закономерности предметной области.

К такого типа формальным языкам относятся, в частности, SADT (Structured Analysis and Design Technique – методология структурного анализа и проектирования) и возникшее на основе SADT семейство языков IDEF, широко используемое разработчиками программного обеспечения в описании и управлении бизнесом, и вообще во всех случаях работы со сложными системами. Большим достоинством этих языков является то, что они основаны на использовании графики, что позволяет человеку, с одной стороны, легко излагать свои

представления, в том числе и образные, а с другой, понимать чужие.

К такому типу разработок следует, по-видимому, отнести онтологию проектирования, развиваемую Н.М. Боргестом и коллегами.

Достоинством идеологии, основанной на использовании формальных языков (в случае их удачной разработки) является возможность описания ситуации с минимальными затратами, и последующего эффективного управления ситуацией.

Недостаток, - затруднения или даже невозможность нахождения эффективного решения в случае, когда ситуация не описывается первоначальными упрощающими представлениями. В таком случае решение ищется человеком путем выхода за пределы первоначальной модели и другими методами. И для поддержки интеллектуальных усилий человека в этих случаях необходимо разрабатывать онтологии второго типа.

2) Разработка формальных семантик, в максимальной (на современном уровне) степени учитывающей особенности понимания и мышления человека. Такая семантика должна разрабатываться в комплексе с:

- а) базой знаний, содержащих как знания о мире, так и знания о семантике в целом, и семантике информационных единиц различных уровней;
- б) интеллектом, способным генерировать новые смыслы. Возможно, для этого недостаточно использования символьных методов, и нужно привлекать иные методы, в частности, нейросети.

В контексте необходимости более мощной семантики невозможно не упомянуть работы Д.А. Поспелова [Поспелов, 1986] и его последователей, которые на несколько десятилетий опередили время. Их представления о семантике как модели ситуации очень актуальны и в наше время, и, конечно, их нужно использовать и развивать.

Предлагаемое направление разработок весьма трудозатратно. Возникает вопрос: а где же брать деньги?

Представляется, что для этого необходимо организовать бизнес-проект в области Semantic Web, объединяющий усилия специалистов из Белоруссии, России и Украины. Ежегодный рынок услуг в этой области составляет многие миллиарды. Необходимо создание собственной поисковой машины, основанной на семантике нового поколения, и раскрутка сервера под эту поисковую машину, сравнимого по масштабу с Yahoo, Altavista, Google...

Заключение

В работе рассмотрены особенности реализации семантики в искусственных системах и у человека. Рассмотрены ключевые компоненты, составляющие

семантику. Показано, что лишь незначительная и очень простая их часть реализуется в искусственном интеллекте. Наиболее важным недостатком формальной семантики является ее неспособность к построению высших уровней смысла сообщения, а именно на этом уровне понимается информация, происходит мышление и принимаются решения во всех случаях, когда ситуация отличается от типовой.

На основании анализа литературных данных о сущности онтологий и областей их применения дано определение онтологии как инвариантного, наиболее экономного и последовательного представления содержания некоторой предметной области или вида деятельности, основанного на определенной модели семантики, а также одном или нескольких стандартизированных языках представления данных.

Разработаны требования к моделям семантики и онтологиям, основанные прежде всего на использовании принципов понимания информации человеком. Сделан вывод, что реализации семантики нового типа необходимо совместно реализовывать и использовать интеллект, онтологию и базы знаний о мире и о семантике.

Предложено реализовать разработку новой семантики для поиска информации как бизнес-проект в рамках развития Semantic Web.

Библиографический список

- [Berners-Lee et al., 2001] Berners-Lee T., Hendler J., Lassila O. The Semantic Web // Scientific American -2001. – V.284, N 5. – P.29-37.
- [Colomb, 2007] Colomb R.M. Ontology and the Semantic Web – IOS Press – 2007. – 272 p.
- [Guarino et al., 2009] Guarino N., Oberle D., Staab S. What Is an Ontology / Eds. Staab S., Studer R. – Springer-Verlag – Berlin – 2009. – p.1-17.
- [Studer et al., 1998] Studer R., Benjamins R., Fensel D. Knowledge engineering: Principles and methods // Data and Knowledge Engineering – 1998/ - V.25, N 1-2. – P.161-198.
- [Shadbolt et al., 2006] Shadbolt N., Hall W., Berners-Lee T. The Semantic Web Revisited // IEEE Intelligent Systems – 2006. – V.21, Issue 3 – P.96-101.
- [Shepherd, 2004] The Synaptic Organization of the Brain – Ed. Shepherd G.M. – Oxford – University Press – 2004. – 736 p. (P.6)
- [Айвазян и др., 1998] Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики – М.: ЮНИТИ – 1998. – 1022 с. (с.565).
- [Васильев, 1990] Васильев Л.М. Современная лингвистическая семантика: Учеб. пособие для вузов – М.: Высш. шк. – 1990. – 176 с.
- [Голенков и др., 2012] Голенков В.В., Гулякина Н.А. Графодинамические модели параллельной обработки знаний: Принципы построения, реализации и проектирования // Минск – БГУИР – 2012. – С.23-52.
- [Кобозева, 2000] Кобозева И.М. Лингвистическая семантика: Уч. пособие – М.: Эдиториал УРСС. – 2000. – 352 с.
- [Кудрявцев, 2010] Кудрявцев Д.В. Системы управления знаниями и применение онтологий: Учеб. пособие – СПб – Изд-во Политехн. ун-та – 2010. – 344 с.
- [Кузнецов, 2012] Кузнецов О.П. Когнитивная семантика: новые подходы к старым проблемам // XII Международная научная конференция имени Т.А. Таран «Интеллектуальный анализ информации» - Киев – 2012 – С.77-88.
- [Ландэ, 2005] Ландэ Д. Семантический веб: от идеи – к технологии // Телеком – 2005. - № 6.
- [Ланцош, 1965] Ланцош К. Вариационные принципы механики – М.: Мир – 1965. – 406 с.

[Нариньяни, 2001] Нариньяни А.С. Кентавр по имени Теон: тезаурус + онтология / / Международный семинар по компьютерной лингвистике и ее приложениям «Диалог 2001» - 2001. – Т.1. – С.184-188.

[Питерс, 1999] Питерс Т. Аутизм: От теоретического понимания к педагогическому воздействию – СПб.: Институт специальной педагогики и психологии – 1999. – 192 с. (с.29-30)

[Попова и др., 2007] Попова З.Д., Стернин И.А. Когнитивная лингвистика – М.: АСТ: Восток-Запад – 2007. – 314 с.

[Поспелов, 1986] Поспелов Д.А. Ситуационное управление: теория и практика – М.: Гл. ред. физ.-мат. лит. – 1986. – 288 с.

[Поспелов, 1989] Поспелов Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов – М.: Радио и связь – 1989. – 184 с.

[Сторож, 2013] Сторож В.В. Актуальные направления исследований мышления и интеллекта // Искусственный интеллект – 2013. - № 3. – С.30-37.

[Шведин, 2011] Шведин Б.Я. Онтология проектирования – Terra Incognita? // Онтология проектирования – 2011. - № 1. – С.9-21.

[Ярцева, 1989] Большой энциклопедический словарь. Языкознание – Под ред. В.Н. Ярцевой – Советская энциклопедия – 1989. – 687 с.

SEMANTICS AND ONTOLOGY IN NATURAL AND ARTIFICIAL SYSTEMS

Storozh V.V. *

*Donetsk State Technical University

ws52@mail.ru

In work features of realisation of semantics in artificial systems and at the human are considered. The key components making semantics are considered. It is shown that only insignificant and their very simple part is realised in artificial intelligence. The most important lack of formal semantics is its inability to construction of the highest levels of sense of the message. However at this level the information is understood, there is a thinking and decisions in all cases when the situation differs from the typical are made.

On the basis of the analysis of the literary data definition ontology as invariant, most economical and consecutive representation of the maintenance of some subject domain or activity kind is made. Thus ontology as a basis should use certain model of semantics, and also one or several standardised languages of data presentation.

Requirements to models of semantics and ontology are developed. These requirements are based first of all on use of principles of understanding of the information by the human. The conclusion is drawn that it is necessary for realisation of semantics of new type to realise and use in common intelligence, ontology and knowledge bases about the world and about semantics.

It is offered to realise working out of new semantics for information search as the business project within the limits of development Semantic Web.