МОДЕЛИРОВАНИЕ МЫШЛЕНИЯ И ФОРМАЛИЗАЦИЯ РЕФЛЕКСИИ

II. ОНТОЛОГИИ И ФОРМАЛИЗАЦИЯ ПОНЯТИЙ*

Д.Е. Пальчунов

Введение

Статья посвящена исследованию методов формализации человеческого мышления для задач моделирования интеллектуальной деятельности человека на компьютере. Центральным является следующий вопрос: в каком объеме творческая, исследовательская деятельность человека может быть формализована и насколько она может быть реализована на вычислительной машине? Есть ли здесь какие-либо принципиальные ограничения или это только вопрос техники и времени? Или по-настоящему творческая деятельность человека принципиально не может быть выполнена компьютером?

Указанную проблему можно поставить и в несколько ином виде: может ли быть формализована и реализована на компьютере исследовательская деятельность научных сообществ [1])? С одной стороны, такая формулировка проблемы является значительно более сложной, чем первая. Действительно, одно дело промоделировать деятельность одного человека, другое – тысяч людей, причем не только деятельность их самих, но и многоуровневые взаимодействия внутри человеческого общества. Но с другой стороны, если ограничиться задачей формализации только исследовательской деятельности, вторая проблема может в определенных аспектах оказаться проще первой.

---

* Работа поддержана грантами Междисциплинарного интеграционного проекта СО РАН № 1 «Вычислимость и рациональность: исследование сферы применимости тезиса Черча – Тьюринга и понятия эффективного вычисления к проблеме соотношения дедуктивного и эмпирического способов познания когнитивных и физических процессов» и РФФИ № 05-01-04003-НИИО-а (DFG project COMO, GZ: 436 RUS 113/829/0-1).

Существует много подходов к точной постановке и решению проблемы возможности реализации на компьютере интеллектуальной деятельности человека. В частности, решение этой проблемы связывается с теоремой Геделя о неполноте, а сама проблема сводится к существованию разрешающего или перечисляющего алгоритма для математических теорем. Например, известно, что множество тождественно истинных утверждений арифметики перечислимо, но неразрешимо (как следует из теоремы Геделя).

Однако с нашей точки зрения, формулировка проблемы реализации на компьютере творческой деятельности человека в таком виде, как «существует ли программа машины Тьюринга, которая будет перечислять тексты (всех) содержательных математических теорем (более точно, номера этих текстов)?», не является удовлетворительной.

Сформулируем кратко наши основные аргументы против такой формулировки.

1. Если мы потребуем перечислимости множества номеров *всех* содержательных математических теорем, то такое требование будет слишком сильным. Действительно, почему должен существовать единый алгоритм, описывающий мышление всех математиков (живших, живущих и будущих жить впоследствии)? Мы можем допустить существование такого алгоритма для каждого математика в отдельности, но при этом предположить, что множество номеров всех этих алгоритмов не является перечислимым.

Таким образом, требование перечислимости множества номеров всех содержательных математических теорем оказывается слишком сильным. С другой стороны, даже требование разрешимости этого множества является чрезвычайно слабым. Поясним это более подробно.

2. Предположим, что математика будет успешно развиваться еще миллиард лет, а после этого вся Вселенная исчезнет (например, взорвется). Очевидно, что в таком случае множество всех полученных математических теорем будет конечным и, следовательно, разрешимым. Поэтому существует и программа машины Тьюринга, которая породит в точности номера всех этих теорем. Также очевидно и то, что наличие такой программы совершенно не дает ответа на вопрос «Может ли компьютер мыслить?».

Данный аргумент можно сформулировать и не прибегая к предположению о столь трагичном завершении математики. Сформулируем содержательную проблему так: «Может ли компьютер получить формулировки всех теорем, которые математики докажут в ближайшие 100 (1 000, 1 000 000) лет?». Эта проблема, безусловно, является весьма интересной, однако если ее переформулировать в таком виде: «существует ли программа машины Тьюринга, порождающая в точности формулировки всех теорем, которые математики докажут в ближайшие 100 (1 000, 1 000 000) лет?», – то ответ очевидным образом будет положительным в силу конечности этих множеств теорем.

3. И наконец, указанный аргумент можно сформулировать еще одним способом. Очевидно, что формулировка содержательной математической теоремы не может иметь длину, превышающую $10^{10^{10^{10}}}$ символов. Множество таких формулировок конечно и опять же разрешимо.

Поэтому, с одной стороны, требование пречислимости множества номеров содержательных математических теорем является слишком сильным, а с другой стороны, даже требование разрешимости этого множества является чрезвычайно слабым.

Легко видеть причину неудовлетворительности таких требований: пусть программа существует, но где ее взять? Каков будет объем такой программы – не будет ли она размером со Вселенную? Проще ли написать текст этой программы, чем получить все математические теоремы, которые она будет порождать?

По существу, мы имеем дело с проблемой получения новых знаний, и вопрос состоит в том, может ли их получать компьютер. Здесь необходимо понять, что такое новое знание. Например, с нашей точки зрения, новая математическая теорема является новым знанием, в то время как с точки зрения Л. Витгенштейна, знание, содержащееся в теореме, уже содержится в аксиомах и потому не является новым.

Очевидно, что если формулировки (и доказательства) теорем уже «зашиты» в программу, то компьютер, работающий по такой программе и порождающий эти формулировки, новых знаний не производит. Поэтому, с нашей точки зрения, вопрос состоит вовсе не в том, существует ли программа, перечисляющая формулировки теорем. Проблема заключается в том, существует ли обозримый набор формальных принципов и правил, такой что исходя из них компьютер сможет порождать новые знания, в явном виде в этих принципах не содержащиеся и, соответственно, неизвестные человеку, который их сформулировал. Если такой набор принципов существует, то компьютер может осуществлять творческую интеллектуальную деятельность. И наоборот, если компьютер сможет доказывать новые содержательные математические теоремы, то программу этого компьютера можно в определенном смысле рассматривать как указанный набор принципов и правил.

Такой набор принципов и правил можно пытаться получить с помощью формализации и моделирования интеллектуальной, творческой деятельности человека. С нашей точки зрения, одним из наиболее важных моментов такой деятельности являются порождение научных понятий и оперирование ими. В частности, мы уже отмечали, что формулировка содержательной математической теоремы должна быть обозримой. Заметим, что использование новых понятий может принципиально сокращать формулировки теорем, краткость текста формулировки связана с наличием «сильных» понятий. Цель данной статьи – исследовать возможность формализации порождения и использования научных понятий.

В своей интеллектуальной деятельности человек использует различные понятия, причем организованные в сложную иерархию. В нашем исследовании эту иерархию мы называем онтологией или иерархией онтологий. Исследователь за свою жизнь придумывает несколько новых понятий, наиболее талантливые исследователи – несколько десятков. Однако любой исследователь в своей деятельности пользуется сотнями тысяч понятий. Как он овладевает этими понятиями? В процессе обучения. Каким образом это происходит? Как организован процесс обучения? Можно ли этот процесс формализовать?

Изучению того, как у ребенка формируются понятия, посвящено очень много работ. В наибольшей степени в этом продвинулись Л.С. Выготский [2] и Ж. Пеаже [3]. Однако не только нет сколько-нибудь окончательного решения этой проблемы, но и не видно близких перспектив ее решения.

Отметим, что обучающиеся системы, которые сегодня реализованы на компьютере, работают, по существу, не с понятиями, а с низкоуровневыми данными. По большей части такие системы основаны на нейронных сетях. Таким образом, в настоящее время не реализовано обучение программных систем понятиям, а также не формализован процесс порождения новых понятий.

Следовательно, для решения задачи формализации и компьютерного моделирования творческой деятельности исследователя мы должны формализовать как процесс возникновения новых понятий, так и процесс обучения уже известным понятиям.

С другой стороны, для формализации исследовательской деятельности научного сообщества нужно решить только одну из этих проблем – формализовать процесс возникновения новых понятий. Действительно, если рассматривать научное сообщество в целом и за весь период его развития, то можно считать, что оно только порождает новые понятия, но не обучается им, поскольку все научные понятия, которые оно использует, оно же и произвело. Вот почему достаточно формализовать только процесс возникновения новых понятий. В этом смысле задача моделирования исследовательской деятельности научного сообщества проще, чем задача моделирования творческой деятельности одного исследователя.

Один из возможных способов решения проблемы использования понятий, возникающей при моделировании исследовательской деятельности человека, может быть следующим: не обучать программную систему новым понятиям, а вместо этого все известные понятия сразу в нее вложить. При этом, однако, возникают две сложности. Во-первых, для того чтобы успешно работать, исследователю постоянно приходится знакомиться с новыми понятиями, введенными другими исследователями, т.е. обучаться этим новым понятиям. Во-вторых, необходимо переходить от одних узких предметных областей к другим, при этом смысл понятий будет меняться, т.е. опять же придется, по существу, изучать новые понятия. Поэтому полностью решить проблему данным способом вряд ли удастся. Тем не менее в нашем исследовании мы будем использовать и эту возможность – заранее задавать иерархию известных понятий, представленную как иерархия онтологий предметных областей.

В настоящей работе исследуются различные формальные подходы к определению понятий. В качестве одного из подходов к алгебраическому представлению смысла понятий рассматривается анализ формальных понятий [4], в котором исследуются формальные контексты и решетки формальных понятий. Мы рассматриваем сильные и слабые стороны этого подхода. Продолжена разработка теоретико-модельного подхода к формализации онтологий предметных областей, начатая нами ранее [5]. Введена графовая структура на множестве аксиом теории первого порядка и на ее сигнатуре – множестве ключевых понятий предметной области. На основе сети понятий введены представления о степени близости между различными понятиями и расстоянии между понятиями. Теоретико-модельную формализацию онтологий предметных областей мы используем для формального определения глоссария.

Математическая формализация понятий

Как было указано выше, исключительно важной проблемой является формальное описание процессов возникновения новых понятий и овладения неизвестными понятиями. Мы используем аппарат теории моделей для разработки подхода к решению этой проблемы.

Здесь необходимо отметить следующее. Овладение новыми понятиями естественно подразделить на два принципиально разных случая: осознание понятий ребенком и постижение новых понятий взрослым. Например, ребенку необходимо осознать, что такое собака. Этот процесс достаточно сложный, протяженный во времени и состоит из нескольких этапов [6]. Взрослый также сталкивается с новыми понятиями, например узнает, что среди пород собак есть бультерьер. Однако у взрослого овладение новым понятием «бультерьер» имеет достаточно мало общего с тем, как ребенок осознает, что такое собака. Как минимум, взрослый строит для себя явное, вербальное описание свойств бультерьера, которое для него, по существу, является определением этого понятия (причем его «определение», очевидно, не будет полностью совпадать с каким-либо общепринятым определением, даваемым экспертами по собаководству). Ребенок же формирует понятие собаки, основанное на образах.

Еще один важный момент, который следует отметить, состоит в том, что понятия в сознании человека претерпевают изменения, причем порой очень существенные. Например, человек, живущий в деревне и хорошо знакомый с понятием «фрукт», будет поражен, познакомившись с разнообразными тропическими фруктами. При этом его понятие фрукта существенно изменится.

Таким образом, мы должны формализовать не только овладение новыми понятиями, но и развитие понятий в сознании человека в процессе его познавательной деятельности.

Следующее важное разделение видов понятий, которые мы должны формализовать, это разделение на обыденные (или «верхнеуровневые») понятия и научные понятия. К первым относятся такие понятия, как «вода», «собака», «животное», «вещество», «человек» и т.д. Это понятия, которые употребляются всеми людьми и имеют некоторый общепринятый смысл. Таким образом, эти понятия принадлежат к наиболее общей, «верхнеуровневой» онтологии. Второй тип понятий – научные понятия, которые являются терминами специальных дисциплин. Это такие понятия, как «треугольник», «алгебраическая система», «интеграл», «электрон», «молекула» и т.д. Безусловно, «верхнеуровневые» понятия могут стать терминами специальных дисциплин, – например, понятие «вода» приобретает в химии вполне определенный и точный смысл, а газированная вода уже водой не является, но является раствором. С другой стороны, многие научные понятия перекочевывают из научной лексики в обыденную, например такие понятия, как «треугольник», «электричество», «лазер».

Тем не менее имеются существенные различия между обыденными и научными понятиями как в закономерностях их возникновения в человеческом обществе, так и в способах овладения этими понятиями отдельным человеком. Имеются также существенные различия между обыденными и научными понятиями в их содержании и полноте определения. Данное разделение понятий на обыденные и научные в определенной степени сходно с разделением понятий на эмпирические и теоретические, которое проводил, например, Р. Карнап [7].

Для формализации понятий нам необходимо использовать математику. В настоящее время в качестве основного математического аппарата формализации понятий рассматривается математическая логика. Однако ряд исследователей подвергают критике математическую логику (в частности, логику предикатов) как единственное основание для математической формализации понятий. Не проводя детального анализа этой критики, мы укажем наиболее существенные, с нашей точки зрения, недостатки математической логики как способа формализации понятий естественного человеческого мышления. Некоторые другие важные недостатки логики предикатов, связанные как с отсутствием структуры аксиом теории, так и с отсутствием структуры множества понятий – сигнатуры, мы рассмотрим несколько позже.

В логике предикатов понятия формализуются в виде одноместных и многоместных предикатов. Одноместные предикаты используются для формализации понятий-свойств, многоместные – для формализации понятий-отношений. При этом, что весьма существенно, предикат отождествляется с его объемом – множеством элементов (для свойства) или кортежей элементов (для n-местного отношения), на которых данный предикат является истинным. Однако такое отождествление, очень удобное с математической точки зрения, не полностью приемлемо с точки зрения содержания.

Рассмотрим, например, понятие «собака». Каждый взрослый человек знает, что такое собака. Более того, даже если с течением времени это понятие и меняется в его сознании, то весьма медленно и незначительно. С другой стороны, в реальном мире объем понятия «собака» изменяется каждую минуту, а может, и каждую секунду: одни собаки рождаются, другие умирают. Кроме того, легко заметить, что человек, прекрасно понимающий, что такое собака, имеет достаточно слабое представление об объеме этого понятия. В частности, большинство людей вряд ли смогут более или менее точно оценить мощность объема этого понятия – количество собак, живущих в данный момент на Земле (а тем более, живших за всю их историю).

Таким образом, принятое в математической логике отождествление предикатов с их объемами является весьма условным. Такое отождествление сходно с моментальной «фотографией» реального мира: если мы рассматриваем только данный момент времени, то отождествление предиката с его объемом вполне допустимо. Но если мы рассмотрим ситуацию в динамике, то увидим принципиальное различие между предикатом и его объемом: предикат остается достаточно неизменным, в то время как его объем постоянно меняется.

Имеются и другие подходы к математической формализации понятий реального человеческого мышления, отличные от представления понятий в логике предикатов. Ниже мы остановимся на некоторых из них подробнее.

Анализ формальных понятий

Одним из наиболее разработанных подходов к логическому анализу порождения и представления понятий является *анализ формальных понятий*, предложенный Рудольфом Вилле [8]. Этот подход основывается на идеях, изложенных в книге А. Арно и П. Николя «Логика, или Искусство мышления», изданной в 1662 г. в Париже и более известной как «Логика Порт-Рояля» [9].

При формальном исследовании понятий Р. Вилле и его последователи отталкиваются от необходимости замены традиционного математического подхода к формализации понятий как n-местных предикатов, основанного на классической математической логике. С нашей точки зрения, критика Вилле математической логики за то, что она не отражает существа естественных понятий, которыми мыслит человек, является достаточно обоснованной. Он и его ученики предложили другой способ математической формализации реальных понятий – на языке формальных контекстов и теории решеток. Эта формализация, получившая название анализа формальных понятий, имеет достаточно много практических приложений. Такие приложения основаны в первую очередь на визуальном представлении набора понятий в виде *решетки формальных понятий*.

Вилле предложил принципиально новый подход к формализации понятий. Вместо рассмотрения предиката как множества истинности, т.е. по существу, как множества элементов или кортежей элементов данного основного множества, понятие представляется как пара: объем и множество признаков – экстент (extent) и интенд (intend). Таким образом, различаются сущность понятия – предиката и его объем – множество его значений истинности. Между содержаниями предикатов и их объемами имеются соотношения, известные в математике как соответствия Галуа.

Как мы уже отмечали выше, одной из наиболее важных и сложных проблем, которые возникают при попытке формального описания творческой, исследовательской деятельности человека, является проблема порождения новых понятий. Эта проблема возникает, например, при автоматизации доказательства теорем. В настоящее время направление исследований в области автоматического доказательства теорем, по существу, свелось к разработке методов создания пруверов, осуществляющих логический вывод в разрешимых фрагментах логики предикатов, таких, например, как логика описаний. Сейчас стало совершенно очевидным, что задача формулировки и доказательства новых (и, что очень важно, интересных) математических теорем далеко не сводится к реализации логического вывода на компьютере. Необходимо формализовать процесс математического творчества, одной из важных составляющих которого является порождение новых математических понятий.

Центральным моментом порождения новых (в том числе и математических) понятий являются осознание того, что определенный набор интересующих нас объектов обладает некоторыми общими свойствами, выделение и точная формулировка этих свойств. Типичный пример этого – введение Клейном понятия группы. Таким образом, в некотором рассматриваемом контексте выделяется набор объектов, обладающих рядом общих признаков. Если объекты являются действительно интересными, а признаки – существенными, вводится новое понятие, включающее в себя указанные признаки; объекты же, обладающие этими признаками, рассматриваются как примеры объема нового понятия.

Такой путь позволяет переходить от элементарных, «низкоуровневых» понятий к более сложным, «высокоуровневым». Анализ формальных понятий дает точную математическую формализацию указанного подхода к порождению новых понятий. Интересен вопрос: является ли такой способ единственным или существуют другие механизмы возникновения новых понятий?

Приведем в кратком изложении основные определения и факты, на которых основывается анализ формальных понятий [10].

Рассматриваются *формальные объекты* и *формальные атрибуты* – признаки формальных объектов. *Формальным контекстом* называется тройка: множество формальных объектов, множество формальных атрибутов и *отношение инцидентности* – соответствие между множеством объектов и множеством атрибутов, означающее, что данный объект удовлетворяет данному атрибуту. Такой способ представления информации является очень распространенным. Например, в анализе данных и теории распознавания образов данные представляются именно в виде таблиц «объекты – признаки» [11].

На формальных контекстах вводятся *формальные понятия*. Формальное понятие является парой, состоящей из двух множеств: это его *объем* – множество формальных объектов и его *содержание* – множество формальных атрибутов. При этом должно быть выполнено два условия: во-первых, каждый объект, принадлежащий к объему формального понятия, должен иметь все атрибуты из его содержания, и, во-вторых, каждый атрибут из содержания формального понятия должен принадлежать ко всем объектам из его объема.

Более того, такая пара множеств должна быть максимальной по отношению включения.

Дадим точные определения.

Определение 1. Тройка множеств (G, M, I) называется *формальным контекстом*, если $I \subseteq G \times M$. При этом множество G называется *множеством формальных объектов*, множество M – *множеством формальных атрибутов*, а соответствие I – *отношением инцидентности*.

Пару (A, B) назовем *формальным понятием* формального контекста (G, M, I), если $A \subseteq G$, $B \subseteq M$, $A = \{g \in G \mid gIm$ для всех $m \in B\}$ и $B = \{m \in M \mid gIm$ для всех $g \in A\}$.

Множество A называется *объемом* (экстент, extent) формального понятия (A, B), а множество B называется *содержанием* (интенд, intend) формального понятия (A, B).

Рассмотрим формальный контекст (G, M, I) и два формальных понятия (A, B) и (C, D) формального контекста (G, M, I). Обозначим $(A, B) \le (C, D)$ если $A \subseteq C$.

Определенное таким образом формальное понятие очень близко к одному из центральных объектов изучения теории моделей – аксиоматизируемому классу. Каждый аксиоматизируемый класс является формальным понятием, если в качестве формальных объектов рассматривать алгебраические системы, в качестве формальных атрибутов – предложения, а в качестве отношения инцидентности – истинность предложения на модели. Взаимосвязь между формальными понятиями и аксиоматизируемыми классами была подробно изучена нами ранее [12].

Одним из важных действий над формальными контекстами и понятиями является порождение формального понятия по множеству объектов и порождение формального понятия по множеству атрибутов. В качестве иллюстрации приведем следующие примеры.

Понятия из атрибутов:

- {красный} порождает понятие красных объектов;
- {красный, фрукт} порождает понятие красных фруктов;
- {красный, фрукт, яблоко} порождает понятие красных яблок.

Понятия из объектов:

• {Институт математики СО РАН} порождает понятие Института математики СО РАН;

• {Институт математики СО РАН, Институт философии СО РАН} порождает понятие института СО РАН;

• {Институт математики СО РАН, ИАПУ ДВО РАН} порождает понятие института российской академии наук.

Замечание 1. Множество $\boldsymbol{B}(\boldsymbol{K}) = \boldsymbol{B}(G, M, I)$ формальных понятий произвольного формального контекста $\boldsymbol{K} = (G, M, I)$ является полной решеткой $(\boldsymbol{B}(\boldsymbol{K}), \leq)$.

Теорема 1 (основная теорема о решетках формальных понятий) [13]**.** Каждая полная решетка изоморфна решетке формальных понятий некоторого формального контекста.

Подход, предложенный группой Р. Вилле, – формальный анализ понятий позволил существенно продвинуться в решении проблемы формального описания того, как возникают понятия. Однако с нашей точки зрения, этот подход не является окончательным решением проблемы математической формализации понятий и требует дальнейшей разработки как формальных, математических методов, так и самой методологии. Наряду с очевидными достоинствами введенное здесь определение формального понятия имеет и некоторые недостатки. Отметим следующие из них.

1. Далеко не все понятия, которые мы употребляем, являются явно определимыми. В первую очередь это относится к обыденным понятиям, которые человек постигает в раннем детстве (в период овладения своим родным языком). Например, никто не сможет в явном виде указать все признаки, которым удовлетворяют все собаки. Более того, утверждение о том, что только собаки удовлетворяют этой совокупности признаков, является неочевидным и, вполне возможно, неверным. Кроме того, множество признаков, которым удовлетворяет совокупность всех собак, постоянно меняется – со смертью каждой собаки и с рождением новой. Например, очевидно, что постоянно меняется максимальный и минимальный вес собак, а вместе с ним меняется и принадлежность к множеству всех атрибутов собак утверждений вида $wd \geq x$, где wd – вес собаки, а x – количество килограммов (выраженное в виде вещественного числа).

Возражение против этого аргумента может быть следующим: есть существенные и несущественные признаки. Например, признак, что вес собаки не превышает 55 кг, несущественно отличается от признака, что вес собаки не превышает 55,001 кг. Но в таком случае мы должны разделить все признаки на существенные и несущественные (например, для блох разница в весе в 1 г является весьма существенной). Такое разделение приводит нас к следующему недостатку анализа формальных понятий.

2. Когда мы говорим о естественных понятиях, которые мы можем определить более или менее явно, речь идет о достаточно кратких определениях. Например, летчик – это человек, который управляет самолетом, водитель – человек, который управляет автомобилем, и т.д. Никто, определяя какое-то понятие, не приводит три десятка признаков. С другой стороны, класс объектов, задаваемых данным кратким определением, может иметь достаточно богатый набор общих признаков (который, как было указано выше, человек, употребляющий данное понятие, может даже и не представлять). Таким образом, мы можем сформулировать необходимое свойство естественного понятия: оно определяется достаточно небольшим набором признаков. Исходя из известного психологического закона мы можем даже в явном виде уточнить смысл этого «небольшого набора признаков»: их должно быть не более чем 7 ± 2. Этот принцип известен как «закон 7 ± 2», открытый Г. Эббингаузом [14]. Данное свойство очень похоже на понятие конечно аксиоматизируемого класса, только бесконечность заменяется на «очень большое», а конечность – на «обозримое», в данном случае не превышающее 7 ± 2. Поясним это сходство более подробно.

Как уже было отмечено выше, определение формального понятия в некотором смысле идентично определению аксиоматизируемого класса, если мы заменим формальные объекты моделями (т.е. алгебраическими системами), а признаки – предложениями соответствующей сигнатуры. Напомним, что класс $\boldsymbol{K}$ алгебраических систем называется аксиоматизируемым, если он является множеством всех моделей, на которых верны все предложения из данного множества Γ, называемого множеством аксиом рассматриваемого класса. Однако, аксиоматизируемый класс может, вообще говоря, задаваться (аксиоматизироваться) бесконечным множеством аксиом. Класс $\boldsymbol{K}$ алгебраических систем называется конечно аксиоматизируемым, если он является множеством всех моделей, на которых верны все предложения из данного *конечного* множества аксиом Г.

Если алгебраические системы (модели) заменить на объекты произвольной природы, а предложения заменить на признаки, то вместо аксиоматизируемого класса мы в точности получим формальное понятие в данном формальном контексте, состоящем из объектов, признаков и отношения инцидентности, определяющего, что данный объект удовлетворяет данному признаку. Если представить себе бесконечный формальный контекст, то реальным понятиям будут соответствовать формальные понятия, которые задаются конечным набором признаков, – полный аналог конечно аксиоматизируемых классов. Однако, с одной стороны, в реальной практике, как правило, возникают конечные формальные контексты: в них оба множества – множество объектов и множество признаков – конечны. С другой стороны, как было отмечено выше, конечности множества признаков, определяющих данное формальное понятие, еще недостаточно, – необходимо, чтобы это множество признаков было достаточно мало, обозримо, например не превышало 7 ± 2 элементов. Поэтому для конечных формальных контекстов аналогом аксиоматизируемого класса является формальное понятие: оно задается произвольным количеством признаков. Аналогом же конечно аксиоматизируемого класса является «естественное» в нашем смысле понятие – формальное понятие, которое задается обозримым, не превышающим 7 ± 2 множеством признаков.

3. При определении естественных понятий мы можем использовать не только атрибуты, но также их булевы комбинации: отрицания, дизъюнкции, конъюнкции и импликации. Например, учащийся – это школьник или студент; член академии – академик или член-корреспондент; научная публикация – тезисы, статья или монография. Как правило, в практических приложениях анализа формальных понятий в качестве признаков берутся атомарные, т.е. неделимые, признаки. В таком случае нельзя выразить все возможные естественные определения.

Если же, напротив, в качестве формальных атрибутов взять все булевы комбинации элементарных признаков, то в качестве решетки формальных понятий мы получим конечную булеву алгебру, что не является интересным ни с теоретической, ни с практической точки зрения.

4. Разные понятия могут иметь один и тот же объем – в определенном контексте. Например, в лекционной аудитории объемы понятий «молодежь» и «студенты» совпадают, точно так же как совпадают объемы понятий «работник университета» и «преподаватель». Однако это совсем не означает, что даже в этом контексте данные понятия совпадают. Здесь снова имеет место различие между статичным и динамичным представлениями: в настоящий момент множества истинности двух понятий могут совпадать, но это совершенно не означает, что множества истинности этих понятий совпадут в следующий момент.

5. Формальные понятия в формальном контексте, по существу, задают только одноместные отношения. Однако понятия могут определять отношения любой местности, т.е. с любым количеством аргументов.

Казалось бы, эту проблему можно решить, если в качестве формальных объектов брать не только реальные объекты, но также их пары, тройки и т.д., т.е. конечные кортежи реальных объектов. В формальном анализе понятий рассматриваются так называемые *степенные последовательности* – последовательности формальных контекстов, формальными объектами которых являются пары, тройки и т.д. исходных объектов, а формальными атрибутами – соответственно одноместные, двуместные, трехместные и т.д. отношения. Однако такой способ представления является весьма громоздким: вместо одного рассматривается много формальных контекстов и соответствующих им решеток формальных понятий. Кроме того, если для представления отношений разной местности мы рассматриваем различные формальные контексты, то нет возможности выразить связь между отношениями разной местности. Такую, например, как «мать $(x, y) \Leftrightarrow$ родитель (x, y) & женщина (x)».

Можно формальные контексты, соответствующие предикатам разной местности, объединить в один формальный контекст, взяв объединение всех формальных объектов, объединение всех формальных атрибутов и объединение всех отношений инцидентности. Но тогда получившаяся решетка формальных понятий будет исключительно громоздкой. Кроме того, нужно будет вводить дополнительные искусственные предикаты, чтобы отличать объекты от пар объектов, троек объектов и т.д. При таком загромождении формального контекста он потеряет свою наглядность, а вместе с тем и свою полезность.

6. Имеет смысл разделять «признаковые понятия», т.е. понятия, определяемые некоторым набором признаков, и «объектные понятия», которые задаются набором типичных примеров объектов (или кортежей объектов в случае понятия, являющегося многоместным отношением). Признаковые понятия, как уже отмечалось выше, должны определяться небольшим набором наиболее существенных признаков – состоящим из не более чем 7 ± 2 элементов. То же самое должно быть выполнено и для объектных понятий: они должны задаваться одним примером или небольшим количеством примеров (как говорят французы, пример поясняет, примеры запутывают).

Таким образом, анализ формальных понятий является серьезным продвижением на пути формализации того, как возникают новые понятия в реальном человеческом мышлении. Тем не менее результаты, полученные в рамках этого направления, еще очень далеки от окончательного решения проблемы математической формализации понятий.

Два подхода к исследованию онтологий

Выше мы показали исключительную важность формализации реальных понятий, с помощью которых осуществляется человеческое мышление. Формализация понятий является одной из ключевых задач для моделирования на компьютере исследовательской деятельности человека.

В настоящее время одним из наиболее мощных средств представления понятий являются онтологии. Они широко применяются как в теоретических исследованиях, так и в практических приложениях. Приведем основные идеи нашего подхода к логической формализации онтологий предметных областей.

В инженерии знаний онтология описывает терминологическую базу данной предметной области или данного класса предметных областей. Онтология представляет собой формальное описание предметной области, цель которого – в явном виде определить смысл терминов, специфичных для данной предметной области. Наиболее кратко эту функцию онтологий определил Т.Р. Грюбер [15]: «онтология – это явная спецификация концептуализации».

В настоящее время имеется огромное количество исследований по онтологиям [16]. Их можно условно разделить на два направления:

1) исследование методов построения онтологий конкретных предметных областей; применение онтологий для решения различных задач, связанных в первую очередь с моделированием предметных областей [17];

2) разработка логических языков и программных средств для создания и использования онтологий, развитие методов обработки информации, представленной с помощью этих языков [18].

Первое из указанных направлений исследований по онтологиям, по существу, является частью инженерии знаний. Для специалистов, занимающихся этим направлением, онтология есть инструмент моделирования реальности, онтология описывает определенную предметную область. Для точного и гарантированно верного описания некоторой предметной области необходимо сначала по возможности полно специфицировать смысл ключевых понятий, терминов, на языке которых специалисты говорят о данной предметной области. Поэтому онтология должна содержать глоссарий ключевых понятий, дающий спецификацию их смысла. Правильность онтологии определяется тем, что все эксперты в данной предметной области должны признавать утверждения, представленные в онтологии этой предметной области; таким образом, знание, содержащееся в онтологии, должно быть интерсубъективным.

Отметим, что в огромном количестве работ по онтологиям, относящихся к данному направлению инженерии знаний, указывалась необходимость переиспользования знаний, представленных в онтологии. Поэтому одним из наиболее важных, центральных свойств онтологии предметной области является то, что она должна описывать общие свойства предметной области, не зависящие от ее конкретных реализаций.

Второе направление исследований по онтологиям возникло значительно позже первого. В достаточно большой степени это направление связано с деятельностью W3C – WWW-Консорциума, в частности с его проектом «Semantic Web» [19].

В рамках этого направления основное внимание уделяется средствам представления онтологий: логическим языкам, языкам формализации онтологий, программным средствам, ориентированным на разработку онтологий. Например, WWW-Консорциум в рамках проекта «Semantic Web» продвигает язык представления онтологий OWL, основанный на семействе логик Description Logics [20]. Разработка онтологий на языке OWL поддерживается программной системой Protégé.

При этом практически никакого внимания не уделяется тому, какая информация должна быть представлена в онтологии. Возникает впечатление, что с точки зрения представителей этого направления, онтология – это любая информация о предметной области, написанная на языке OWL. Характерно, что в одной из основных работ, лежащих в русле этого направления, – «Справочной книге по онтологиям» [21], по существу, не дается определения онтологии. В качестве определения онтологии (единственного в этой справочной книге!) приводится все то же краткое, символическое определение Грюбера: «онтология – это явная спецификация концептуализации».

В настоящее время в рамках этого направления создано достаточно большое количество инструментальных средств разработки и представления онтологий: Protégé, Protégé-OWL, OilEd, Ontolingua, OntoEdit, WebOnto, OntoSaurus, ODE и т.д. Все они основаны на различных логических средствах.

Возникает вопрос: какими должны и могут быть логические средства представления онтологий? Каков спектр таких средств – от самых простых до самых сложных и выразительных? Возможны следующие варианты ответов на эти вопросы:

- это может быть набор простых отношений между понятиями, например очень популярные в объектно ориентированном программировании отношения «общее – частное», «часть – целое» и т.п;
- в качестве основы языка представления онтологий может быть использована логика описаний (Description Logic);
- и вообще, для представления онтологий может быть взят произвольный разрешимый фрагмент логики предикатов (например, универсальные формулы, формулы с ограниченными кванторами и т.п.).

Какими в таком случае должны быть требования к логическим средствам представления онтологий? В качестве основных можно выделить следующие требования:

• разрешимость – существование алгоритмов проверки доказуемости и эквивалентности формул, непротиворечивости конечных множеств формул;

• эффективность разрешающих алгоритмов;

• наличие уже разработанных пруверов (программ, осуществляющих логический вывод).

Таким образом, следует отметить, что указанные исследования по инструментальным средствам представления онтологий, несмотря на их достаточно большую практическую значимость, по существу, не имеют никакого отношения к содержанию онтологии, т.е. к информации о предметной области, которая в ней представлена.

С нашей точки зрения, напротив, важно, какая именно содержательная информация представлена в онтологии. Способ же реализации онтологии зависит от конкретных задач, для решения которых мы используем онтологии предметных областей.

Таким образом, при исследовании и разработке онтологий мы имеем дело с двумя основными проблемами:

• что такое онтология, т.е. какую информацию о предметной области она должна содержать;

• каким способом должна быть представлена онтология для эффективной работы с ней (в частности, необходимы разрешимость и не очень большая сложность разрешающих алгоритмов).

Последняя проблема, как было отмечено выше, достаточно успешно решается в рамках второго направления в исследовании онтологий. Мы же сконцентрируем внимание на первой проблеме. Ее мы можем более детально сформулировать в виде следующих проблем:

• что такое онтология?

• чем отличается онтология от не-онтологии?

• какую информацию о предметной области должна содержать онтология?

Исходя из этого в контексте приведенного выше сравнения двух подходов к исследованию онтологий сформулируем основные вопросы, касающиеся определения и формализации онтологий:

1) должна ли (может ли) онтология содержать всю известную информацию о предметной области?

2) верно ли, что любая информация о предметной области, записанная на языке OWL, может быть включена в онтологию данной предметной области?

3) достаточна ли логика описаний (Description Logic) как логическое средство представления онтологий?

4) насколько широко может применяться автоматизация разработки онтологий, каковы методы такой автоматизации?

5) насколько эффективным может быть применение онтологий, какие практические задачи, связанные с моделированием предметных областей, могут быть решены посредством применения онтологий?

Подход к формализации онтологий, развиваемый в настоящей работе, позволяет ответить на первые два вопроса. Отметим сразу, что наш ответ на оба этих вопроса будет отрицательным.

Кратко напомним основные идеи логической формализации онтологий, изложенной в первой части данной работы [22].

Для нашего подхода к определению онтологии наиболее важно то, какую информацию о предметной области она содержит. Основной вопрос, который мы при этом решаем, – какие сведения о предметной области должна содержать онтология, чтобы ее можно было гарантированно переиспользовать. Это означает, что информация, представленная в онтологии, должна быть верна для любого экземпляра описываемой предметной области.

Для того чтобы дать формальное определение онтологии предметной области [23], мы применяем подход Р. Карнапа к классификации истинности различных высказываний [24]. Карнап рассмотрел три типа истинности предложений: логическую истинность, аналитическую истинность и синтетическую истинность. Утверждение является логическим, если значение истинности этого утверждения зависит только от его логической формы. Таким образом, предложение $(\varphi \vee \neg\varphi)$ является логически истинным, а предложение $(\varphi \& \neg\varphi)$ – логически ложным.

Утверждение является аналитическим, если его значение истинности зависит только от смысла понятий, которые используются в данном утверждении. Поэтому предложение «У вдовы когда-то был муж» – аналитически истинное, а «Корова является птицей» – аналитически ложное.

Высказывание является синтетическим, если его значение истинности определяется состоянием реального мира. Например, предложение «Земля вращается вокруг Солнца» является синтетически истинным, в то время как предложение «Земля – центр Вселенной» – синтетически ложным.

Как мы уже отметили, онтология должна описывать общие свойства предметной области, не зависящие ее конкретной реализации, – это дает возможность ее повторного использования. Таким образом, онтология предметной области должна содержать только ту информацию, которая является верной для каждого конкретного примера данной предметной области. Поэтому онтология должна содержать только предложения, являющиеся аналитическими в контексте рассматриваемой предметной области [25]. Онтология предметной области должна содержать множество ключевых понятий предметной области и аналитические предложения, определяющие смысл данных ключевых понятий.

Определение 2. Формальной онтологией предметной области *SD* назовем пару $O = <A, \sigma>$, где σ – множество ключевых понятий предметной области, и *A* – множество аналитических предложений, описывающих смысл этих ключевых понятий. Множество *T* предложений, которые являются верными в каждом примере предметной области, будем называть теорией предметной области *SD.*

Теоретико-модельное представление онтологии как средство формализации понятий

В этом разделе мы продолжим разработку методов логической формализации онтологий предметных областей, которую проводили в первой части данной работы [26]. Теоретико-модельный подход к представлению онтологий мы применим для формализации понятий естественного языка.

Для дальнейшего изложения напомним определения и обозначения по теории моделей [27]. Набор $\sigma < P1, \ldots, Pn, f1, \ldots, fk, c1, \ldots, cm >$, где $P1, \ldots, Pn$ – символы предикатов, $f1, \ldots, fk$ – символы функций и $c1, \ldots, cm$ – символы констант называется сигнатурой. $S(\sigma)$ обозначает множество всех предложений сигнатуры σ; $\sigma(\varphi)$ обозначает сигнатуру предложения φ, то есть множество всех сигнатурных символов, входящих в φ; $\sigma(\Gamma)$ обозначает сигнатуру множества предложений Г. Мы пишем $\Gamma + \psi$ если из множества предложений Г выводимо предложение ψ.

Напомним, что теорией называется дедуктивно замкнутое множество предложений. Пусть T – теория и $\sigma = \sigma(T)$. Если для множества предложений $\Gamma \subseteq S(\sigma)$ выполнено $T = \{\psi \in S(\sigma) \mid \Gamma + \psi\}$, то мы говорим, что Г является множеством аксиом теории T.

Для множества предложений Г множество $Th(\Gamma) = \{\psi \in S(\sigma(\Gamma)) \mid \Gamma + \psi\}$ называется теорией, аксиоматизируемой множеством предложений Г; в частности, $Th(\varphi) = Th(\{\varphi\})$ – это теория, аксиоматизируемая предложением φ. Символ $\subset$ обозначает строгое включение: для множеств A и B имеем $A \subset B$ если $A \subseteq B$ и $A \neq B$.

Заметим, что в определении 2 онтологии предметной области $O = <A, \sigma>$, фактически, множество σ – это сигнатура онтологии. Это означает, что σ содержит только символы понятий. Множество A включает определения символов, содержащихся в сигнатуре σ.

Кроме того, выполнено $\sigma \subseteq \sigma(A)$, но не обязательно верно $\sigma = \sigma(A)$. Это означает, что множество аналитических предложений A может содержать сигнатурные символы, которые не являются символами ключевых понятий предметной области. Такое происходит тогда, когда при описании смысла сигнатурных символов (т.е. символов ключевых понятий) мы используем утверждения, содержащие понятия, не являющиеся ключевыми.

Для формальной онтологии $O = <A, \sigma>$ множество A в общем случае не является теорией, т.е. не является дедуктивно замкнутым.

Определение 3. Пусть пара $<A, \sigma>$ является формальной онтологией предметной области *SD* и пусть $\sigma_0 = \sigma(A)$. Множество $T_a = \{\varphi \in S(\sigma_0) \mid A + \varphi\}$ будем называть аналитической теорией предметной области *SD*.

Таким образом, на языке логики предикатов первого порядка мы определили, что такое онтология предметной области, теория предметной области и аналитическая теория предметной области.

Логика предикатов как средство описания понятий реального человеческого мышления имеет один существенный недостаток. Теория первого порядка является просто множеством, т.е. представляет собой своеобразный «мешок» предложений, которые никак не упорядочены и не имеют никакой внутренней структуры. Сигнатура теории является таким же множеством, «мешком» понятий, также не имеющим никакой структуры. Более того, любое конечное множество предложений можно «склеить» в одно предложение – взяв их конъюнкцию.

Таким образом, теория первого порядка (которая, очевидно, всегда бесконечна) может быть представлена как нарастающая последовательность гигантских предложений, каждое из которых включает в себя все предыдущие (в виде конъюнктивных членов). Очевидно, что такая ситуация неадекватна закономерностям реального мышления, которое имеет дело только с достаточно маленькими и обозримыми высказываниями.

Поэтому при формализации понятий нашей первой задачей будет придание структуры множествам предложений логики предикатов первого порядка. Для решения этой задачи мы будем использовать теоретико-модельный подход к онтологиям.

Рассмотрим типичный пример фрагмента онтологии (рис. 1). Здесь фрагмент онтологии представлен в виде размеченного графа, в узлах которого стоят понятия (термины предметной области); граф является ориентированным, а его ребра помечены – в данном случае отношениями «часть – целое» и «общее – частное» (наиболее популярными отношениями в объектно ориентированном программировании и возникшем под его влиянием подходе к формализации онтологий, реализованном, например, в OWL [28]).

Наряду со своими очевидными достоинствами – простотой и наглядностью такой способ представления онтологий имеет и ряд существенных недостатков. Отметим два наиболее важных из них.

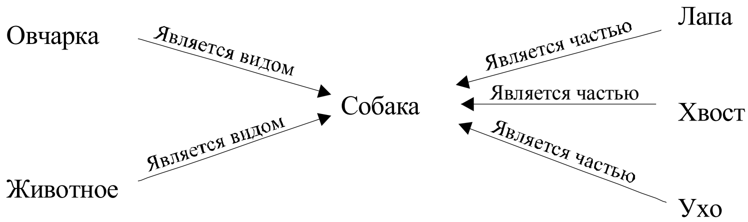

Рис. 1

Во-первых, при таком способе представления онтологий мы можем формализовать только двуместные отношения. Однако на практике часто необходимо иметь дело с многоместными отношениями, которые нельзя свести к двуместным. Типичным примером из математики является трехместное отношение компланарности векторов. Из обыденной жизни можно привести другой простой пример: три человека x, y и z встречались вместе. Не вводя дополнительных сущностей, это отношение нельзя представить как композицию двуместных отношений.

Во-вторых, использование в графе помеченных ребер приводит к необходимости зафиксировать рассматриваемый набор двуместных отношений. А это является весьма неудобным, когда возникает необходимость расширять и дополнять онтологию, поскольку могут появляться новые отношения. Таким образом, добавление в онтологию новых понятий – двуместных отношений будет неизбежно приводить к необходимости перестройки всего графа, а именно, изменения набора меток для его ребер.

Из этой ситуации есть один очень простой выход – представление онтологии в виде лингвистической сети [29] (рис. 2). В данном случае мы имеем несколько иное устройство графа. У него вершины двух типов: понятия и предложения. На рисунке 2 понятия изображены в виде овалов, а предложения – в виде прямоугольников. Граф не ориентирован и имеет только один вид ребер: каждое ребро соеди-

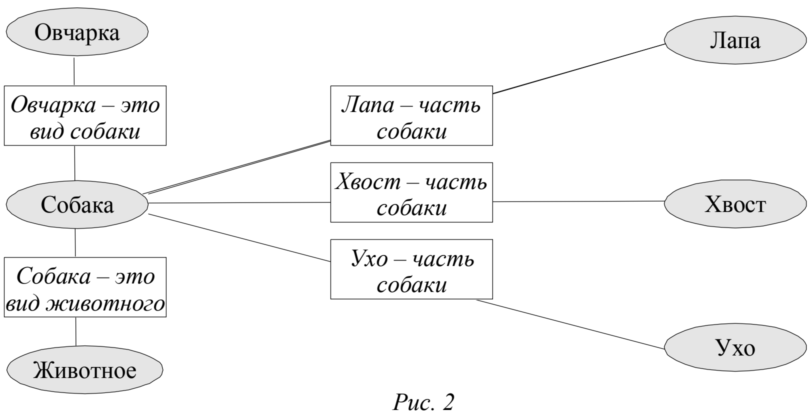

Рис. 2

няет понятие и предложение, причем в том случае, если данное понятие входит в данное предложение.

Таким образом, мы получаем диаграмму представления онтологии в виде лингвистической сети (рис. 3). Такое представление очень популярно в направлении инженерии знаний, которое имеет дело с автоматизацией извлечения информации из текстов естественного языка [30]. Оно дает нам ключ к структурному представлению предложений логики предикатов.

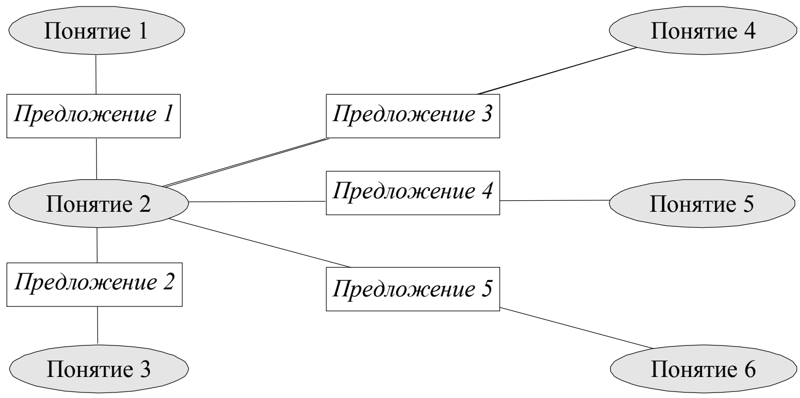

Рис. 3

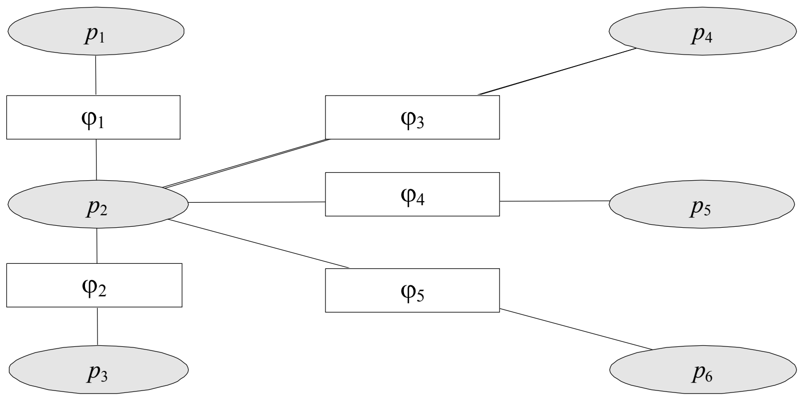

Рис. 4

Рассмотрим прямое обобщение лингвистической сети – синтаксическую сеть предложений логики предикатов. На рисунке 4 представлен неориентированный граф, который имеет вершины двух типов: сигнатурные символы и предложения логики предикатов первого порядка. Ребра графа соединяют понятия с предложениями, в сигнатуру которых они входят.

Дадим точное определение формальной синтаксической сети.

Определение 4. Пусть σ – сигнатура, $\Gamma \subseteq S(\sigma)$ и $\sigma = \sigma(\Gamma)$. Тройку $<\Gamma, \sigma, R>$ назовем *формальной синтаксической сетью,* если $R \subseteq \Gamma \times \sigma$ и для любых $\varphi \in \Gamma$ и $p \in \sigma$ выполнено $(\varphi, p) \in R$ тогда и только тогда, когда $p \in \sigma(\varphi)$.

Недостатком формальной синтаксической сети является то, что сигнатурные символы могут входить в предложение формально, но не по существу. Например, если предложение φ не содержит сигнатурного символа p, мы можем рассмотреть эквивалентное ему предложение $\varphi\&(\psi\vee\neg\psi)\equiv\varphi$, содержащее p (взяв предложение ψ, которое содержит сигнатурный символ p). В таком случае предложение φ не связано в графе с сигнатурным символом p, а эквивалентное ему предложение ψ связано с p, что плохо с семантической точки зрения. Для решения этой проблемы мы введем понятие формальной семантической сети.

Определение 5. Пусть σ – сигнатура, $\Gamma \subseteq S(\sigma)$ и $\sigma = \sigma(\Gamma)$. Формальную синтаксическую сеть $<\Gamma, \sigma, R>$ назовем *формальной семантической сетью,* если для любых $\varphi\in\Gamma$ и $p\in\sigma$ из $(\varphi, p) \in R$ следует, что для каждого $\psi \equiv \varphi$ выполнено $p \in \sigma(\psi)$.

Из теоремы Крейга [31] вытекает

Теорема 2 [32]**.** Для каждого предложения первого порядка можно найти эквивалентное ему предложение, содержащее наименьшее по включению множество сигнатурных символов. Т.е. для каждого предложения φ существует предложение $\psi \equiv \varphi$, такое что для любого предложения $\xi \equiv \psi$ выполнено $\sigma(\psi) \subseteq \sigma(\xi)$.

Используя теорему 2, мы показываем, что имеет место

Теорема 3. Для любого множества предложений Γ существует формальная семантическая сеть $<\Gamma_0, \sigma, R>$, такая что $\{[\varphi]_{\equiv} \mid \varphi \in \Gamma\} = \{[\varphi]_{\equiv} \mid \varphi \in \Gamma_0\}$.

Это означает, что каждое множество предложений может быть представлено формальной семантической сетью, совпадающей с исходным множеством предложений с точностью до эквивалентности формул.

Здесь через $[\varphi]_{\equiv}$ мы обозначаем множество предложений, эквивалентных предложению φ.

Таким образом, любую теорию первого порядка мы можем корректно представить в виде формальной семантической сети.

Следствие 1. Пусть T – теория сигнатуры σ. Для теории T существует формальная семантическая сеть $<\Gamma, \sigma, R>$, которая ее представляет, т.е. выполнено $\{[\varphi]_{\equiv} \mid \varphi \in T\} = \{[\varphi]_{\equiv} \mid \varphi \in \Gamma\}$.

Следующая проблема состоит в том, что при представлении теории в виде семантической сети любое конечное количество узлов сети будет «слипаться» в один: $\{\varphi_1, \ldots, \varphi_n\} \to \varphi_1 \,\&\, \ldots \,\&\, \varphi_n$. Как мы уже отмечали выше, в настоящий момент эта проблема является весьма существенным недостатком представления информации на языке теорий логики предикатов первого порядка. Для ее решения мы введем понятие канонического множества аксиом теории.

Определение 6. Пусть T – теория, $\sigma = \sigma(T)$, $\Gamma \subseteq S(\sigma)$ и Γ являются множеством аксиом для теории T, т.е. $T = \{\psi \in S(\sigma) \mid \Gamma + \psi\}$. Назовем Γ *каноническим множеством аксиом* теории T если выполнено:

1) для любых $\varphi \in \Gamma$ и ψ если $\varphi \equiv \psi$, то $\sigma(\varphi) \subseteq \sigma(\psi)$;

2) для любого $\varphi \in \Gamma$ и любых предложений ψ и ξ если $\varphi \equiv \psi \& \xi$, то $\sigma(\psi) \not\subset \sigma(\varphi)$ либо $\sigma(\xi) \not\subset \sigma(\varphi)$.

Пункт (2) определения 6 означает, что не существует предложений $\varphi \in \Gamma$, ψ и ξ, таких что $\varphi \equiv \psi \& \xi$, $\sigma(\psi) \subset \sigma(\varphi)$ и $\sigma(\xi) \subset \sigma(\varphi)$. В противном случае вместо φ в множество аксиом Γ следовало бы включить предложения ψ и ξ, содержащие меньшее количество сигнатурных символов.

Таким образом, если множество предложений Γ является каноническим множеством аксиом некоторой теории, то оно не содержит «склеенных» предложений; каждое содержащееся в нем предложение минимально с точки зрения его сигнатуры. Ответом на естественный вопрос, всякую ли теорию можно представить таким образом, является следующее утверждение.

Теорема 4. Для любой теории существует каноническое множество ее аксиом.

Достаточно очевидным является

Следствие 2. Для любой теории существует наибольшее по включению каноническое множество аксиом.

Однако ответ на следующий вопрос не очевиден.

Проблема 1. Существует ли для произвольной теории минимальное по включению каноническое множество аксиом?

Эта проблема имеет положительное решение для очень важного на практике частного случая.

Следствие 3. Конечно аксиоматизируемая теория имеет минимальное по включению каноническое множество аксиом.

Предложение 1. Существует теория, которая не имеет множества аксиом Γ, такого что множество $\{[\varphi]_{\equiv} \mid \varphi \in \Gamma\}$ было бы наименьшим по включению.

В связи с этим возникает аналогичный вопрос для конечно аксиоматизируемых теорий.

Проблема 2. Верно ли, что всякая конечно аксиоматизируемая теория имеет каноническое множество аксиом Γ, такое что множество $\{[\varphi]_{\equiv} \mid \varphi\in\Gamma\}$ является наименьшим по включению? Верно ли, что всякая конечно аксиоматизируемая теория имеет минимальное по включению каноническое множество аксиом Γ, такое что множество $\{[\varphi]_{\equiv} \mid \varphi\in\Gamma\}$ является наименьшим по включению?

Мы уже отмечали выше, что недостатком теорий первого порядка как средства представления знаний является то, что понятия (как и предложения) не имеют никакой внутренней структуры. Сигнатура – это просто множество символов, никак не структурированных и не упорядоченных. Данный недостаток не имеет никакого значения в математике, поскольку в ней, как правило, рассматриваются объекты очень маленькой сигнатуры. Например, сигнатура групп состоит из трех символов: двух операций – умножения и взятия обратного элемента и одной константы – единицы. Сигнатура колец состоит из пяти элементов, сигнатура булевых алгебр – также из пяти элементов, а сигнатура упорядоченных множеств состоит вообще из одного элемента – двуместного отношения порядка.

Однако ситуация кардинальным образом меняется, когда мы начинаем рассматривать модели реальных систем, например бизнес-модель предприятия. Сигнатура такой модели может содержать тысячи, а то и десятки и сотни тысяч понятий: номенклатуру товаров, сырья, оборудования и т.д. В таком случае структура множества сигнатурных понятий становится чрезвычайно важной. В частности, очень полезно определять степень близости понятий, относятся ли они к одной ситуации (одному производственному процессу) или, напротив, между ними практически нет ничего общего. Мы предлагаем решение этой проблемы, основанное на придании структуры множеству сигнатурных символов элементарной теории логики предикатов первого порядка.

Основываясь на представленит теории при помощи канонического множества аксиом, мы введем сеть понятий. А именно, из семантической сети предложений логики предикатов (см. рис. 4), в качестве множества предложений которой взято каноническое множество аксиом некоторой теории, мы получаем сеть понятий, в узлах которой стоят понятия – сигнатурные символы, причем ребро соединяет два понятия, если они входят в сигнатуру одного предложения (рис. 5). На первой картинке (см. рис. 4) мы попросту «стерли» предложения и получили вторую картинку (см. рис. 5). Из семантической сети предложений мы получили сеть понятий.

Дадим формальное описание указанных структур.

***Определение** 7*. Пусть $\Gamma \subseteq S(\sigma)$, $\sigma = \sigma(\Gamma)$, пусть $<\Gamma, \sigma, R>$ – формальная семантическая сеть. Пару $<\sigma, Q>$ назовем *сетью понятий*,

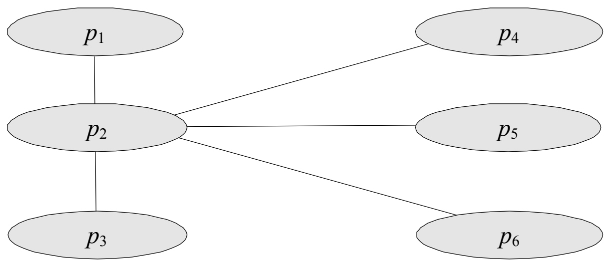

Рис. 5

порожденной формальной семантической сетью <Γ, σ, *R*> если для любых p, $q \in \sigma$ выполнено: $(p, q) \in Q$ тогда и только тогда, когда найдется предложение $\varphi \in \Gamma$, для которого $(\varphi, p) \in R$ и $(\varphi, q) \in \mathrm{R}$.

Таким образом, теперь мы имеем естественную структуру на множестве понятий. Два понятия максимально близки, если они входят в одну аксиому. Это означает, что они описывают один процесс, или одну сущность, или одну роль в производственном процессе и т.п.

Более того, теперь мы можем говорить о степени близости понятий, подсчитывать расстояние между понятиями – длину минимального пути от одного понятия до другого.

Заметим, что данное определение степени близости понятий действительно будет осмысленным, если в формальной семантической сети <Γ, σ, *R*> множество Γ будет каноническим множеством аксиом (некоторой теории). Если же, напротив, в качестве Γ взять некоторую теорию, то, как легко увидеть, все понятия будут связанными и поэтому расстояние между любыми двумя понятиями будет равно единице. В таком случае данное определение близости понятий становится бессмысленным.

Важно отметить, что степень близости понятий зависит от теории, в контексте которой мы их рассматриваем. Это не только напрямую следует из наших определений (т.е. имеет место для нашего способа формализации понятий), но и является принципиальным для описания реального мира. Например, понятия «человек» и «планета Марс» в обычной ситуации очень далеки друг от друга. Но когда речь идет о пилотируемой экспедиции на Марс, эти они становятся непосредственно связанными и, соответственно, максимально близкими.

Таким образом, в данном разделе мы предложили способ придания определенной структуры знаниям и понятиям, представленным на языке логики предикатов первого порядка. Для этого мы определили каноническое множество аксиом теории, семантическую сеть предложений и сеть понятий.

Теоретико-модельное представление глоссария

Далее мы применим теоретико-модельные методы для разработки подхода к формализации того, как возникают новые понятия, и того, как человек ими овладевает. В этом разделе мы дадим логическую формализацию глоссария.

Мы уже отмечали выше, что имеется принципиальное различие между тем, как понятия впервые возникают у ребенка, и тем, как взрослый человек узнает новые понятия [33]. Основа этого различия состоит в том, что при овладении новыми понятиями взрослый использует язык, т.е. сводит новые понятия к уже имеющимся. Ребенок же конструирует в своем мозгу понятия «с нуля», на основе только «первой сигнальной системы» – зрительных, слуховых и кинестетических (обонятельных и осязательных) образов.

Механизм образования понятий в мозгу ребенка очень сложен и до конца не изучен. Поэтому мы начнем исследование и формализацию с того, как взрослый человек узнает новые понятия. Типичная ситуация такого познания – овладение специальными понятиями новой, еще не изученной области. Наиболее распространенный способ изложения смысла понятий, специфичных для данной предметной области, – глоссарий (или тезаурус; здесь мы не будем сосредоточивать внимание на различии между глоссарием и тезаурусом).

Главная цель данного раздела – дать на языке теории моделей точное формальное определение глоссария предметной области и исследовать возможность представления смысла понятий при помощи глоссария.

Как правило, глоссарий состоит из статей, в которых дается определение ключевых понятий некоторой предметной области. Часто глоссарием называют просто словарь, который объясняет специфические слова и выражения из какой-либо области знания. Глоссарий (тезаурус) специфических терминов предметной области является наиболее простым способом представления онтологии данной предметной области. Более точно, глоссарий представляет наиболее простую часть онтологии предметной области.

Дадим формальное определение глоссария на языке теории моделей.

Определение 8. Пусть σ – сигнатура. Последовательность предложений $<\varphi_1, \ldots, \varphi_n>$ назовем *формальным глоссарием (определяющим понятия из* σ*)*, если выполнено:

1) $\sigma(\varphi_1) \subset \sigma(\varphi_1 \ \& \ \varphi_2) \subset \ldots \subset \sigma(\varphi_1 \ \& \ \ldots \ \& \ \varphi_n) = \sigma$;

2) добавление каждого нового предложения φ_{k+1} консервативно расширяет предыдущий набор предложений $\varphi_1, \ldots, \varphi_k$, т.е. $Th(\varphi_1 \ \& \ \ldots \ \& \ \varphi_k) = Th(\varphi_1 \ \& \ \ldots \ \& \ \varphi_n) \cap S(\sigma(\varphi_1 \ \& \ \ldots \ \& \ \varphi_k))$.

Консервативность расширений наборов предложений является исключительно важной, поскольку при определении новых терминов мы не должны изменять смысл уже определенных понятий (иначе эти предыдущие «определения» не были определениями в строгом смысле).

В контексте формального определения глоссария полезно ввести новое онтологическое отношение между понятиями: понятие p необходимо для определения понятия q. В частности, понятие p должно быть определено в глоссарии раньше понятия q. Это отношение между понятиями практически не отражено в литературе, хотя оно почти столь же фундаментальное, как и отношения «часть – целое» и «общее – частное». Правда, в отличие от этих отношений отношение «должно быть определено раньше» не является абсолютным: для некоторых отношений оно может зависеть от способа представления глоссария.

Действительно, для некоторых пар понятий p и q можно сначала определить понятие p, а затем понятие q, а можно и наоборот – сначала понятие q, а затем понятие p. Например, можно сначала определить, что такое бегемот, а потом – что такое гиппопотам, а можно наоборот. То же самое верно и для других пар синонимов: «базар» и «рынок» и т.д. Эта ситуация имеет место и для научных терминов: для булевых алгебр и вообще произвольных решеток можно сначала определить операции, а потом через них формульно определить отношение частичного порядка, а можно, наоборот, сначала определить порядок, а затем с его помощью формулами первого порядка определить операции.

Теперь остановимся более подробно на явных и неявных определениях. Явным мы будем называть такое определение, в котором один новый сигнатурный символ явно определяется через предыдущие. Например, определяемое двуместное отношение P выполнено на паре элементов (x, y) тогда и только тогда, когда на этой паре выполнено некоторое утверждение $\psi(x, y)$, содержащее только уже определенные понятия.

Дадим строгое определение явного глоссария.

Определение 9. Формальный глоссарий $<\varphi_1, \ldots, \varphi_n>$ *явно определяет* понятия из σ, если существуют формулы $\psi_1(x), \ldots, \psi_n(x)$ такие, что для каждого $k < n$ выполнено одно из трех следующих условий:

1) $\varphi_{k+1} = \forall x(\mathrm{P}(x) \leftrightarrow \psi_{k+1}(x))$;
2) $\varphi_{k+1} = \forall x \forall y((f(x) = y) \leftrightarrow \psi_{k+1}(x, y))$;
3) $\varphi_{k+1} = \forall y((c = y) \leftrightarrow \psi_{k+1}(y))$,

где $P, f, c \in \sigma \setminus \sigma(\varphi_1 \,\&\, \ldots \,\&\, \varphi_k)$, x – кортеж переменных и $\sigma(\psi_{k+1}) \subseteq \sigma(\varphi_1 \,\&\, \ldots \,\&\, \varphi_k)$.

Альтернативой явных определений являются неявные. В неявном определении формула φ_k содержит один или несколько определяемых сигнатурных символов; она представляет известную нам информацию о смысле этих терминов (понятий), но не дает явного выражения смысла этих понятий через предыдущие, как это имело место в определении 9. Для дальнейшего изложения важно отметить, что в случае явного определения (как в определении 9) предложение φ_k содержит только один новый (т.е. определяемый) сигнатурный символ, в то время как в случае неявного определения предложение φ_k может содержать сколь угодно много новых сигнатурных символов.

Возникает вопрос: а нужны ли неявные определения или, напротив, всегда можно обойтись явными? Всегда ли смысл набора новых понятий предметной области можно задать в виде последовательности явных определений, т.е. явного глоссария?

Заметим, что, как следует из определения 9, в явном глоссарии определяется сначала одно понятие, потом второе, затем третье и т.д. То есть каждый раз новое определение содержит только один новый сигнатурный символ.

Поэтому ослабленной версией предыдущего вопроса является следующий: всегда ли можно построить определение новых понятий по одному, так чтобы каждое определение (пусть даже и неявное) содержало только одно новое понятие?

С помощью следующей теоремы мы дадим отрицательный ответ на оба этих вопроса. Будем говорить, что формальный глоссарий $<\varphi_1, \ldots, \varphi_n>$ представляет предложение ψ, если $Th(\psi) = Th(\varphi_1 \ \& \ \ldots \ \& \ \varphi_n)$.

Теорема 5. Существуют сигнатура σ, состоящая из имен двух понятий, и предложение ψ, определяющее смысл понятий из σ, для которых нет формального глоссария $<\varphi_1, \varphi_2>$, представляющего предложение ψ, такого чтобы сигнатура $\sigma(\varphi_1)$ состояла из имени одного понятия, а сигнатура $\sigma(\varphi_2)$ – из имени другого понятия.

Следствие 4. Не всегда определения понятий могут быть представлены в виде глоссария, определяющего понятия по одному.

Следствие 5. Не всегда определения понятий могут быть представлены в виде явного глоссария.

Заключение

В данной работе мы продолжили изучение вопроса о том, может ли творческая, исследовательская деятельность человека быть формализована и реализована на вычислительной машине. Для решения этого вопроса мы начали исследование и формализацию системы понятий, на которых основано человеческое мышление. В настоящее время одним из наиболее разработанных способов представления понятий являются онтологии. Мы рассмотрели два основных развиваемых в настоящее время подхода к формализации и исследованию онтологий. В рамках первого преимущественное внимание уделяется информации о предметной области, которая представлена в онтологии. В рамках второго подхода, напротив, в центр внимания ставятся логические, языковые и программные средства разработки и представления онтологий. Придерживаясь первого подхода, мы концентрируем наш интерес на содержании онтологии.

В данной работе нас более всего интересует проблема возникновения понятий и овладения понятиями человеком. Пока мы оставляем без рассмотрения более сложный вопрос о формировании понятий у ребенка и исследуем то, как взрослый человек осознает новые понятия. Для взрослого человека одним из основных средств знакомства с новыми, неизвестными ранее понятиями является глоссарий (или тезаурус) предметной области. Поэтому мы исследуем вопрос о формализации глоссария некоторой области знаний. В работе предложено формальное определение глоссария, основанное на языке логики предикатов и теории моделей. Исследованы различные свойства глоссариев: возможность явного определения понятий, возможность определения понятий одного за другим. Показано, что в общем случае смысл понятий не может быть представлен глоссарием, дающим явные определения понятий; точно так же в общем случае является невозможным и определение понятий по одному. Это обусловливает необходимость исследования неявных определений понятий.

* * *

Автор выражает признательность участникам семинара под руководством Ю.Л. Ершова и В.В. Целищева, а также Р. Вилле, К.Э. Вольфу и С.О. Кузнецову за интересные и полезные обсуждения.

Примечания

1. См., например: *Макаров В.Л.* Искусственные общества // Искусственные общества. – 2006. – Т. 1, № 1. – С. 10–24; *Он же.* Компьютерная модель общества // Материалы к Круглому столу «Искусственные миры в экономике». – Воронеж, 2006. – С. 2–6.

2. См.: *Выготский Л.С.* Мышление и речь. – 5-е изд., испр. – М.: Лабиринт, 1999; *Он же.* Педагогическая психология. – 3-е изд. – М., 1996; *Он же.* Мысль и слово // Психология развития как феномен культуры. – Москва; Воронеж, 1996.

3. См.: *Пеаже Ж.* Речь и мышление ребенка. – М.: Педагогика-Пресс, 1994; *Он же.* Как дети образуют математические понятия // Вопросы психологии. – 1966. – № 4. – С. 121–126; *Он же.* Избранные психологические труды. Психология интеллекта. Генезис числа у ребенка. Логика и психология. – М., 1994; *Он же.* Суждение и рассуждение ребенка. – СПб.: СОЮЗ, 1997.

4. См.: *Ganter B., Wille R.* Formal concept analysis: mathematical foundations. – Berlin; Heidelberg: Springer-Verlag, 1999; *Ganter B., Stumme G., Wille R.* Formal concept analysis: foundations and applications. – Berlin; Heidelberg: Springer-Verlag, 2005.

5. См.: *Пальчунов Д.Е.* Моделирование мышления и формализация рефлексии I: Теоретико-модельная формализация онтологии и рефлексии // Философия науки. – 2006. – № 4 (31). – С. 86–14; *Пальчунов Д.Е., Сидорова Е.С.* Виртуальный каталог // Тр. Всерос. конф. «Знания – Онтологии – Теории». – Новосибирск, 2007. – С. 166–175; *Pal'chunov D. E.* GABEK for ontology generation // Learning and Development in Organizations: V. 2 / Ed. by Herdina P., Oberprantacher A., Zelger J. GABEK – Contributions to Knowledge Organization. – Wien: LIT, 2007. – P. 87–107.

6. См.: *Выготский Л.С.* Мышление и речь; *Пеаже Ж.* Речь и мышление ребенка; *Он же.* Как дети образуют математические понятия.

7. См.: *Carnap R*. Philosophical foundations of physics. – New York; London: Basic Books, 1968.

8. См.: *Ganter B., Wille R*. Formal concept analysis: mathematical foundations; *Ganter B., Stumme G., Wille R*. Formal concept analysis: foundations and applications.

9. См.: *Arnauld A., Nicole P*. La logique ou l'art de penser: contenant, outre les regles communes, plusieurs observations nouvelles, propres a former le iugement. – P., 1662.

10. См.: *Ganter B., Wille R*. Formal concept analysis: mathematical foundations; *Ganter B., Stumme G., Wille R*. Formal concept analysis: foundations and applications.

11. См.: *Загоруйко Н.Г*. Методы распознавания и их применение. – М.: Сов. радио, 1972; *Он же*. Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во Ин-та математики СО РАН, 1999.

12. См.: *Palchunov D.E*. Lattices of relatively axiomatizable classes // ICFCA 2007: Lecture Notes in Artificial Intelligence, № 4390 / Ed. by S.O. Kuznetsov and S. Schmidt. – Berlin; Heidelberg: Springer-Verlag, 2007. – P. 221–239.

13. См.: *Ganter B., Wille R*. Formal concept analysis: mathematical foundations.

14. См.: Ebbinghaus H. Uber das Gedachtnis: Untersuchungen zur experimentellen Psychologie. Neue, unverand. und ungek. Ausgabe nach der 1. Aufl. 1885. – Darmstadt: Wissenschaftliche Buchgesellschaft, 1992; *Величковский Б.М*. Современная когнитивная психология. – М.: Изд-во МГУ, 1982.

15. См.: *Gruber T.R*. A translation approach to portable ontologies // Knowledge Acquisition. – 1993. – V. 5, No. 2. – P. 199–220.

16. См.: *Gruber T.R*. A translation approach to portable ontologies; *Id*. Towards principles for the design of ontologies used for knowledge sharing // International Journal of Human-Computer Studies. – 1995. – V. 43, Is. 5–6. – P. 907–928; *Gruber T.R., Olsen G.R*. An ontology for engineering mathematics // Fourth International Conference on Principles of Knowledge Representation and Reasoning / Ed. by J. Doyle, P. Torasso, Sandewall E.; Gustav Stresemann Institut. – Bonn: Morgan Kaufmann, 1994; *Gangemi A., Pisanelli D.M., Steve G*. An overview on the ONIONS Project: Applying ontologies to the integration of medical terminologies // Data and Knowledge Engineering. – 1999. – V. 31, No. 2. – P. 183–220; *Guarino N*. Formal ontology and information systems // Proceedings of International Conference on Formal Ontology in Information Systems (FOIS'98), Trento, Italy / Ed. by N. Guarino. – Amsterdam: IOS Press, 1998. – P. 3–15; *Inaba A., Mizoguchi R*. Learning design palette: An ontology-aware authoring system for learning design: Proc. of International Conference on Computers in Education (ICCE-2004), Melbourne, Australia, Nov. – 30–Dec. 3; *Mizoguchi R*. Ontological engineering: Foundation of the next generation knowledge processing. – WI2001, LNAI2198 / Ed. by N. Zhong et al. Springer-Verlag, 2001. – P. 44–57; *Wielinga B.J., Schreiber A.Th*. Reusable and sharable knowledge bases: A European Perspective // Proceedings KB&KS'93, International Conference on Building and Sharing of Very Large-Scale Knowledge Bases'93, JIPDEC / Ed. by K. Fuchi. – Tokyo, 1993. – P. 103–115; *Клещев А.С., Артемьева И.Л*. Математические модели онтологий предметных областей. Ч. 1–3 // Научно-техническая информация. – Сер. 2 «Информационные процессы и системы». – 2001. – № 2. – С. 20–27; № 3. – С. 19–29; № 4. – С. 10–15; *Fensel D*. OIL: An ontology infrastructure for the semantic Web // IEEE Intelligent Systems. – 2001. – V. 16, No. 2; *Id*. Ontologies: A silver bullet for knowledge management and electronic commerce. – Springer Verlag, 2004; *Maedche A*. Ontology learning for the semantic Web. – Kluwer Academic Publishers, 2002; *OWL* Web ontology language overview, 2003 / Ed. by D. McGuinness., F. Harmelen; *Gomez-Perez A*. Ontology engineering. – Springer-Verlag, 2002; 2003; The handbook on ontologies in information systems / Ed. by S. Staab, R. Studer. – Springer Verlag, 2003.

17. См.: *Gruber T.R*. A translation approach to portable ontologies; *Id*. Towards principles for the design of ontologies used for knowledge sharing // International Journal of Human-Computer Studies. – 1995. – V. 43, Is. 5–6. – P. 907–928; *Gruber T.R., Olsen G.R*. An ontology for engineering mathematics // Fourth International Conference on Principles of Knowledge Representation and Reasoning / Ed. by J. Doyle, P. Torasso, Sandewall E.; Gustav Stresemann Institut. – Bonn: Morgan Kaufmann, 1994; *Gangemi A., Pisanelli D.M., Steve G*. An overview on the ONIONS Project: Applying ontologies to the integration of medical terminologies // Data and Knowledge Engineering. – 1999. – V. 31, No. 2. – P. 183–220; *Guarino N*. Formal ontology and information systems // Proceedings of International Conference on Formal Ontology in Information Systems (FOIS'98), Trento, Italy / Ed. by N. Guarino. – Amsterdam: IOS Press, 1998. – P. 3–15; *Inaba A., Mizoguchi R*. Learning design palette: An ontology-aware authoring system for learning design: Proc. of International Conference on Computers in Education (ICCE-2004), Melbourne, Australia, Nov. – 30–Dec. 3; *Mizoguchi R*. Ontological engineering: Foundation of the next generation knowledge processing. – WI2001, LNAI2198 / Ed. by N. Zhong et al. Springer-Verlag, 2001. – P. 44–57; *Wielinga B.J., Schreiber A.Th*. Reusable and sharable knowledge bases: A European Perspective // Proceedings KB&KS'93, International Conference on Building and Sharing of Very Large-Scale Knowledge Bases'93, JIPDEC / Ed. by K. Fuchi. – Tokyo, 1993. – P. 103–115; *Клещев А.С., Артемьева И.Л.* Математические модели онтологий предметных областей. Ч. 1–3 // Научно-техническая информация. – Сер. 2 «Информационные процессы и системы». – 2001. – № 2. – С. 20–27; № 3. – С. 19–29; № 4. – С. 10–15; *Fensel D*. OIL: An ontology infrastructure for the semantic Web // IEEE Intelligent Systems. – 2001. – V. 16, No. 2; *Id*. Ontologies: A silver bullet for knowledge management and electronic commerce. – Springer Verlag, 2004.

18. См.: *Maedche A*. Ontology learning for the semantic Web. – Kluwer Academic Publishers, 2002; *OWL* Web ontology language overview, 2003 / Ed. by D. McGuinness., F. Harmelen; *Gomez-Perez A*. Ontology engineering. – Springer-Verlag, 2002; 2003; The handbook on ontologies in information systems / Ed. by S. Staab, R. Studer. – Springer Verlag, 2003.

19. См.: *Daconta M.C., Obrst L.J., Smith K.T*. The Semantic Web: A Guide to the future of XML, Web services, and knowledge management. – Wiley Publishing, 2003.

20. См., например: The handbook on ontologies in information systems.

21. Ibid.

22. См.: *Пальчунов Д.Е*. Моделирование мышления и формализация рефлексии I…

23. См.: *Пальчунов Д.Е*. Моделирование мышления и формализация рефлексии I…; *Пальчунов Д.Е., Сидорова Е.С.* Виртуальный каталог; *Pal'chunov D. E*. GABEK for ontology generation.

24. См.: *Carnap R*. Philosophical foundation of physics; *Id*. Meaning and necessity: A study in semantics and modal logic. – Chicago, 1956.

25. См.: *Пальчунов Д.Е*. Моделирование мышления и формализация рефлексии I…; *Pal'chunov D. E*. GABEK for ontology generation.

26. *Пальчунов Д.Е*. Моделирование мышления и формализация рефлексии I…

27. См.: *Ершов Ю.Л., Палютин Е.А.* Математическая логика. – М.: Наука, 1979; *Кейслер Г., Чэн Ч.Ч.* Теория моделей. – М.: Мир, 1977.

28. См.: *OWL* Web ontology language overview, 2003; The handbook on ontologies in information systems.

29. См.: *Zelger J*. GABEK, a new method for qualitative evaluation of interviews and model construction with PC-support // Enchanging human capacity to solve ecological and socio-economic problems / Ed. by E. Stuhler, Suilleabhain M.O. – Munchen; Mering: Rainer

Hampp Verlag, 1993. – P. 128–172; *Id*. Zur Geschichte von GABEK // R. Buber, J. Zelger. GABEK II: Zur Qualitativen Forschung On Qualitative Research / Innsbruck; Wien; Munchen: STUDIENVerlag, 2000. – S. 13–20; *Пальчунов Д.Е*. Синтаксическая близость предложений языка первого порядка // Вычислительные системы. – Вып. 162. Измерение и модели когнитивных процессов. – Новосибирск, 1998. – С. 58–80; *Pal'chunov D.E*. On a logical analysis of GABEK // GABEK II: Zur Qualitativen Forschung On Qualitative Research. – С. 185–203.

30. См.: *Zelger J*. GABEK, a new method…; *Id*. Zur Geschichte von GABEK.

31. См.: *Кейслер Г., Чэн Ч.Ч.* Мышление и речь; *Пиаже Ж*. Избранные психологические труды…

32. См.: *Пальчунов Д.Е*. Алгебраическое описание смысла высказываний естественного языка // Вычислительные системы. – Вып. 158. – Модели когнитивных процессов. – Новосибирск, 1997. – С. 127–148; *Pal'chunov D.E*. Algebraische Beschreibung der Bedeutung von Au?erungen der naturlichen Sprache // GABEK. Verarbeitung und Darstellung von Wissen / J. Zelger, M. Maier. – Innsbruck; Wien: STUDIENVerlag, 1999. – S. 310–326.

Институт математики СО РАН,
Новосибирский государственный университет,
г. Новосибирск

Pal'chunov, D.E. **Simulation of thinking and formalization of reflection: II. Ontologies and formalization of concepts**

The paper studies methods of formalization of human thinking with the view of computer simulation of intellectual activities. Formal definition of a glossary is proposed which is based on the language of predicate logics and model theory. Various glossary properties are considered, viz the ability to define concepts explicitly and the ability to define concepts «one by one». Shown that in general case, the glossary which offers explicit definitions for concepts cannot present their meanings. Similarly, in general case, it is impossible to define concepts «one by one». It makes necessary to study implicit definitions of concepts.