

А.В. Палагин, С.Л. Крытый, Н.Г. Петренко, В.Ю. Величко

Знание-ориентированные информационные системы с обработкой естественно-языковых объектов: онтологический подход

Рассмотрена концепция онтологического подхода к проектированию знание-ориентированных информационных систем с обработкой естественно-языковых объектов. Предложены характерные черты компьютерной системы, спроектированной на основе данного подхода. Приведен пример проектирования компьютерной онтологии предметной области на основе представительного множества естественно-языковых объектов, содержащего около 1000 концептов.

A concept of the ontological approach to designing the knowledge-oriented information systems with the processing of naturally-language objects (NLO) is considered. Characteristic features of the computer system designed on the basis of the ontological approach are suggested. An example of designing the computer ontology of the subject domain on the basis of the representative set of the naturally-language objects containing about 1000 concepts is presented.

Розглянуто концепцію онтологічного підходу до проектування знання-орієнтованих інформаційних систем з обробкою природномовних об'єктів. Запропоновано характерні ознаки комп'ютерної системи, спроектованої на основі даного підходу. Наведено приклад проектування комп'ютерної онтології предметної області на основі репрезентативної множини природномовних об'єктів, що містить близько 1000 концептів.

Введение. В статье описан онтологический подход к разработке методологических и системотехнических основ конструирования и практического применения знание-ориентированных информационных систем (ЗОИС) с обработкой естественно-языковых объектов (ЕЯО). Подход ориентирован на решение проблемы извлечения, формального представления, обработки и системной интеграции знаний. В процессе ее решения выделяют *три аспекта исследования* – онтологический, логический и методологический. Эти аспекты имеют свои, в общем случае фиксированные объекты исследований, соответствующие процессу познания или разработке некоторой ЗОИС. В связи с ними все объекты могут быть разделены на три группы: *система сущностей* (или объектов реального мира), *система знаний* и *система обработки сущностей* в соответствии с данной системой знаний. *Первая* группа представляет собой предмет онтологического исследования (или онтологического подхода), *вторая* – предмет логического исследования и *третья* – методологического. Главными составляющими последнего являются онто-инфологический подход к проектированию класса онтолого-управляемых информационных систем, парадигма двойствен-

ного участия онтологических знаний при языковой и проблемной обработке информации, а также виртуальная парадигма (в частности, архитектура компьютерной системы, ориентированная на технологию реконфигурируемого процессинга) [1]. К онтологическим аспектам относится широкий круг вопросов, начиная с формального описания компонент компьютерных онтологий и заканчивая решением практических задач в конкретных прикладных областях. Главный вектор исследований направлен на формализацию этапов построения, структурирования и представления знаний предметной области (ПдО) и интегрированного с ней проблемного пространства. В свою очередь эффективная реализация указанных этапов и получение конечного результата (в виде базы знаний ПдО) невозможны без проведения системно-онтологического анализа заданной ПдО [2].

Постановка задачи

Прежде чем перейти к описанию сущности онтологического подхода (ОнП) в жизненном цикле ЗОИС (в [1] выделен подкласс онтолого-управляемых информационных систем – ОУИС), обратимся к его философским аспектам, а точнее, к той совокупности общих научных дисциплин, методы и принципы которых непо-

метных областей, и наконец, – к онтологиям процессов и задач [2, 4–12].

Из сказанного следует, что *онтологический подход* к проектированию информационной и архитектурной компонент компьютерной системы (КС) возник как междисциплинарный подход к построению, представлению и обработке структур систем знаний, модели которых описывают структуру и взаимосвязи объектов соответствующих предметных областей.

В статье будут рассмотрены следующие компоненты ОнП¹:

- компьютерные онтологии;
- онтология верхнего уровня;
- онтолого-управляемая архитектура;
- интеграция междисциплинарных знаний и развивающиеся ЗОИС;
- средства поддержки автоматизированного построения О ПдО;
- прикладная лингвистическая направленность ОУИС;
- онто-инфологический подход к проектированию ОУИС;
- режимы функционирования ОУИС.

Онтологический подход

В плане проектирования интеллектуальных информационных систем, ориентированных на решение конкретных задач пользователей, онтологический подход в многочисленной научной литературе трактуется по-разному. Очевидно, это объясняется тем, что не существует его общепринятого определения, а направление онтологического инжиниринга находится в стадии развития и становления. Конкретные сообщества исследователей формулируют только определенные совокупности атрибутов (признаков, свойств, функций), использование которых при проектировании компьютерных систем указывает на ОнП. При этом сам ОнП трактуется как разновидность системного подхода, основанного на знаниях.

Будем полагать, что ЗОИС с обработкой знаний, содержащихся в ЕЯО, спроектирована с

учетом *онтологического подхода*, если она обладает следующими характерными чертами.

• Компьютерные онтологии (КО) обеспечивают эффективную машинную обработку общезыковых и предметных знаний. Такие онтологии названы соответственно [1]: языково-онтологическая картина мира (ЯОКМ) и онтология ПдО. При этом информационным связующим звеном между ними выступает лингвистическая онтология ПдО.

• В отличие от эмпирического подхода – основы «творчества и мудрости», (системно) онтологический подход [2] предполагает строгую (насколько это возможно на текущий момент) систематизацию знаний любого уровня, в том числе категориального. Этот уровень представляется категориальной решеткой, или *онтологией верхнего уровня* (ОВУ) [7, 8, 10]. Проектирование ОВУ должно быть включено в общий алгоритм конструирования ОУИС.

• Архитектурно-технологические особенности:

– онтолого-управляемая архитектура (*ontology-driven architecture – ODA*) характеризуется высоким уровнем формализации представления О ПдО и механизмов онтолого-управляемости (вплоть до развивающихся систем);

– высокая степень системной интеграции междисциплинарных знаний.

• Использование средств поддержки автоматизированного построения онтологий ПдО (методика, технология и программно-аппаратная реализация).

• Прикладная направленность технологии обработки информации и архитектурно-структурной организации ОУИС, поддержанная их сильной взаимосвязью [1].

• Проектирование ОУИС выполняется на основе методов, принципов и механизмов онто-инфологического подхода (ОИП) [1].

• Функционирование ОУИС в двух режимах: накопления предметных онтологических знаний и их обработки.

Архитектурно-структурная организация ОУИС [1] на рис. 2 представлена схемой фрагмента архитектуры, отображающей только ее

¹ Из перечисленных восьми компонент ОнП в данной работе будут рассмотрены первые четыре компонента, а остальные – в следующей публикации.

«онтологическую» компоненту во взаимосвязи с процессами и процедурами (технологической цепочкой) интегрированной информационной технологии (ИИТ) обработки знаний, извлекаемых из ЕЯО. На рис. 2 приняты следующие обозначения: $ОД\ ПдО_k$ – онтология k -го домена предметных областей, где $k = \overline{1, K}$, $K = Card Д$ – мощность множества онтологий доменов, интегрированных в библиотеку; ОИМ – онто-инфологическая модель; ЛО $ПдО_i$ – лингвистическая онтология i -й предметной области, где $i = \overline{1, I}$, $I = Card ПдО$ – мощность множества онтологий предметных областей, интегрированных в библиотеку; 1 – обработка омо-

нимичных форм; 2 – построение поверхностных семантических структур; 3 – обработка семантических неоднозначностей; 4 – обработка концептуальных структур.

Концептуальная модель ИИТ показана на рис. 3.

Компьютерные онтологии – это интенсивно развивающийся раздел информатики как теоретической, так и практической [1, 8, 11]. Актуальность данного направления представляется очевидной в связи с двумя главными обстоятельствами.

Первое из них связано с тем, что КО являются одновременно и результатом развития, и инструментом *knowledge-engineering*, т.е. они вы-

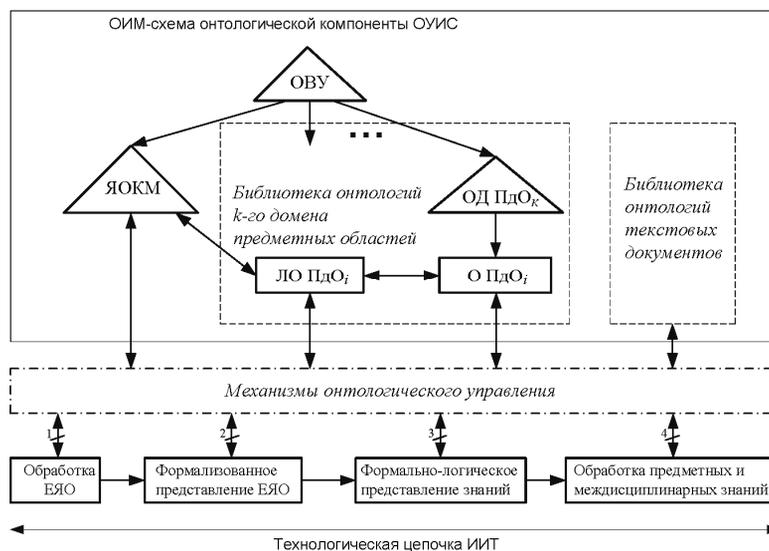


Рис. 2. Фрагмент архитектуры «онтологической» части ОУИС

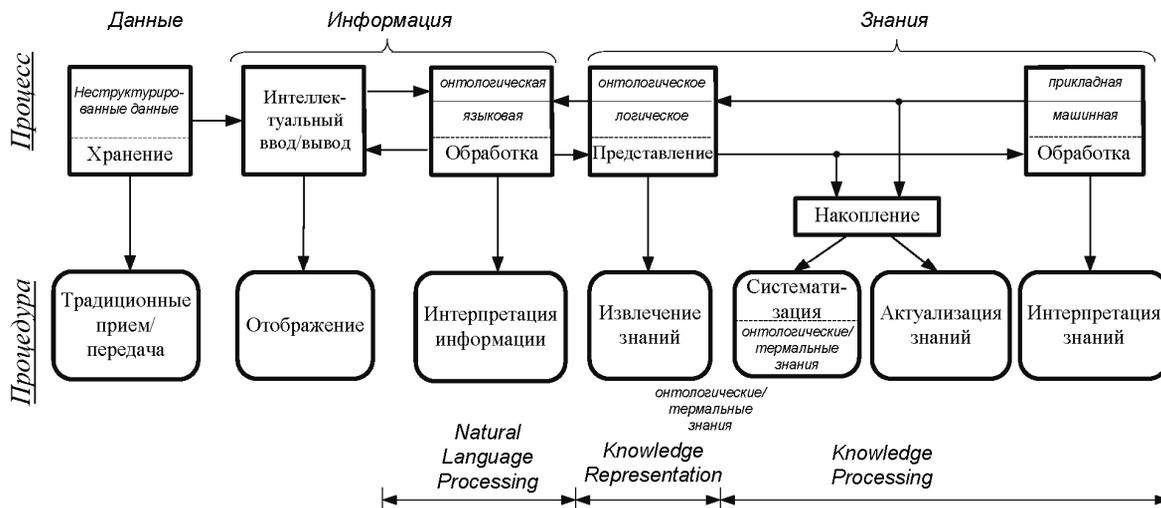


Рис. 3. Концептуальная модель интегрированной информационной технологии обработки знаний, содержащихся в ЕЯО

ступают как способ *концептуализации научной теории*, а также спецификации и формализации БЗ ПдО, выполняя при этом функции классификации, структурирования, упорядочения, интеграции и инструмента при использовании знаний.

Второе обстоятельство связано с функциями онтологий в пространстве современных знаний, которые распространяются в среде Интернет [8]. Подразумевается построение *эффективного механизма поиска информации*, релевантной запросу пользователя, исходя из первичной системы знаний (или начальной естественной онтологии) в интересующей ПдО, и адекватного отображения объекта его интересов в *структурированные семантические модели*, связывающие базовые концепты отношениями частичного порядка, и сложно организованные устойчивые конструкции для формирования запросов и построения ответов с хорошо обозначенной семантикой, а также семантическим маркированием ресурсов и, конечно, с развитием ЗОИС и технологий в широком понимании.

Кроме того, КО некоторой ПдО можно рассматривать как общезначимую, открытую, *формальную научную теорию* (или базу знаний ПдО), представленную на общепринятом (формальном) языке спецификации знаний. Из этого вытекает множество различных практических приложений такой КО, например [9, 10]. Отметим только особый статус такого приложения как обучение. При этом создание «онтологизированных» электронных курсов (в том числе и дистанционных) представляется актуальной самостоятельной, интенсивно развивающейся научно-технической областью.

Понимание терминов онтологического инжиниринга в последнее время существенно прояснилось, и в особенности его «главного» термина – онтологии ПдО, а различения существуют, в основном, в значении и объеме того или иного компонента онтологии [2].

Аксиоматизация научных знаний определяется следующими структурными элементами: аксиомы или схемы аксиом (как не доказываемые в данной системе положения); исходные (первоначальные) понятия (явно не определяемые в системе); правила вывода и построе-

ния системы знаний; определения как правила введения в теорию новых терминов, абстрактных объектов; теоремы как выводимые высказывания; задачи, следствия, положения, выводимые из теории, леммы, логико-методологические принципы аксиоматизирующей теории. Все названные аспекты составляют базовую основу формального научного исследования. Наряду с ними научный подход обязательно предполагает еще одну важную компоненту – интерпретацию. Понятно, что научный смысл имеют лишь те абстракции, которые заведомо могут быть приложены к чему-либо и которые можно сопоставить истине, особенно в тех случаях, когда это практически важно и возможно. Механизмом такого сопоставления истине, поиска логических возможностей ее установления и выступает *интерпретация* как логический прием по установлению значения терминологических выражений теорий, знаковой (формальной) системы [13].

Каждому предложению, логически правильно построенному, в интерпретации придается некоторое значение, например, предложение может быть истинным или ложным. Вопрос же о том, какая из этих возможностей реализована, не решается интерпретацией. Поэтому одна и та же формальная система предполагает множество интерпретаций и только одну истину. Интерпретация гарантирует непротиворечивость теории, но не обеспечивает содержательной истинности данной теории.

Интерпретация имеет свои гносеологические предпосылки. Суть их заключается в том, что в мысленном процессе объективная действительность не просто пассивно отражается, но определенным образом реконструируется, перестраивается. В познании происходит непрерывный процесс образования абстракций и их исключение, перевод одного уровня абстракций на другой. Именно интерпретация и выступает *логическим средством перехода* от одного уровня знаний к другому, средством установления логических связей между этими уровнями [13, 14].

Роль интерпретации отчетливо видна из общей постановки *проблемы извлечения знаний*

из ЕЯО. Пусть T – входной ЕЯТ, в котором «закодированы» отношения $R_i = (R_{i_1}, R_{i_2})$, $i = 1, 2, \dots, n$, где R_{i_1} – синтаксические отношения, а R_{i_2} – семантические отношения. Пусть $X = \{x_1, \dots, x_m\}$ – переменные, означающие словоформы, интерпретированные на T .

Проблема извлечения знаний из ЕЯТ состоит в поиске интерпретации $\varphi: X \rightarrow T$ такой, что $\varphi(X) = \varphi_2(\varphi_1(X))$, где

$$\varphi_1(X) = (\varphi_1(x_1), \dots, \varphi_1(x_m)) \in R_{i_1} \text{ и}$$

$$\varphi_2(\varphi_1(X)) = (\varphi_2(\varphi_1(x_1)), \dots, \varphi_2(\varphi_1(x_m))) \in R_{i_2}.$$

Иными словами, необходимо найти интерпретацию φ – суперпозицию отображений φ_1 и φ_2 , означающие синтаксическую и семантическую интерпретацию, соответственно, предложений текста T . Заметим при этом, что отображение φ_1 несет и некоторую семантическую нагрузку. Она состоит в том, что $\varphi_1(x_i)$, будучи синтаксическим конкретным объектом, имеет естественно-языковое значение. Это дает возможность соотнести $\varphi_1(X) = (\varphi_1(x_1), \dots, \varphi_1(x_m))$ с множеством концептов. Обозначим его через

$$Y = \{(\varphi_1(x_1), \dots, \varphi_1(x_m))\}.$$

Данное, достаточно общее определение проблемы извлечения знаний можно конкретизировать путем уточнения отношений R_i и отображений φ_1 и φ_2 .

Что касается извлеченных знаний, то они представляются в двух видах:

- в виде множества отношений R_{i_2} , определенных на множестве X как база знаний;
- в виде онтологии, отражающей зависимости между концептами $y_j \in Y \subset X$.

В формализованном виде под онтологией понимают упорядоченную тройку $O = \langle Y, R, F \rangle$, где $Y = \{(\varphi_1(x_1), \dots, \varphi_1(x_m))\}$ – множество концептов, $R = \{R_1, R_2, \dots, R_n\}$ – множество отношений на Y и $F = \{\varphi_{21}, \dots, \varphi_{2k}\}$ – множество интерпретаций элементов из $y_j \in Y$ на T . В зави-

симости от конкретизации элементов из Y, R и F различают три типа онтологий [15].

Строгая онтология ПдО (в терминах работы [10] – «весомая» онтология) – это такая онтология, в которой множества концептов и концептуальных отношений максимально полные, а аксиомы представлены определениями и ограничениями. При этом аксиомы представлены на некотором формальном языке, доступном для их интерпретации компьютером.

Более конкретное формальное определение строгой онтологии представляет собой упорядоченную четверку $O = \langle Y, R, F, A(D, Q) \rangle$, где $A(D, Q)$ – множество аксиом, определенных на множестве определений D и отношений Q , ограничивающих концепты, вводимые с помощью определений из D .

Одной из общих конкретизаций (не только строгих онтологий) является отношение частичного порядка на множестве отношений R и интерпретаций F вместе с отношением эквивалентности на множестве концептов Y . Это дает возможность строить иерархию концептов как отдельной онтологии, так и самих онтологий.

Простая онтология ПдО (в терминах работы [10] – «легкая» онтология) – это такая онтология, в которой множество R определено одним типом отношения, а множество F тождественно множеству аксиом ($F \equiv A(D, Q)$). При этом аксиоматизация может быть представлена определениями терминов словаря ПдО на естественном языке (ЕЯ).

Множество промежуточных или *неполных онтологий ПдО* описываются различными сочетаниями мощностей множеств R и F , а аксиоматизация выполнена на ЕЯ.

Один из распространенных вариантов неполной онтологии – структура вида $O = \langle X, R \rangle$, где множество F в явном виде отсутствует ($F = 0$), в предположении, что концепты $y_j \in Y$ общеизвестны (определены по умолчанию) либо (и) достаточно полно интерпретированы отношениями R .

Такая схема классификации по функциональному признаку согласуется с описанием [4]: онтология или концептуальная модель предметной области состоит из иерархии по-

ятий предметной области, связей между ними и законов, которые действуют в рамках этой модели.

В некотором смысле «привилегированной, претендующей на роль метаонтологии» представляется КО – языково-онтологическая картина мира [16]. Это связано, прежде всего, с тем, что языковая форма представления информации – наиболее универсальна, технологический статус языка постоянно возрастает, а сама ЯОКМ может служить предметной онтологией для обработки текстовой информации.

ЯОКМ [17] – открытая, эксплицитно заданная на лексико-смысловом континууме лексикографическая система, в которой совокупность понятий высокого уровня формально обоснована и упорядочена в сложную иерархическую структуру по основным типам лексико-семантических отношений. Различные аспекты конструирования ЯОКМ рассмотрены в [8, 15–17].

Следует отметить, что онтология в реальных ОУИС в общем случае содержит три иерархически связанных компоненты: метаонтологию (онтологию, концептами которой служат онтологии последующих категориальных уровней), предметную онтологию (состоящую из онтологии объектов и онтологии процессов) и онтологию задач.

Онтология верхнего уровня

Термину *компьютерная онтология верхнего уровня* в философской литературе соответствует синоним *система категорий*. Изучению последнего посвящены многочисленные научные труды философов, логиков, лингвистов как зарубежных (древних – от Гераклита до Пирса и Уайтхеда [18], и современных [4–7, 18]), так и советских ученых Апресяна Ю.Д., Сагатовского В.Н., Кузьмина Е.С., Свидерского В.И., Рубашкина В.Ш., Библера В.С., Зиновьева А.А., Соколовской Ж.П. и др. Разработка методологических основ построения многоуровневых компьютерных онтологий предопределила необходимость *системного анализа* указанной научной литературы с целью выявления «рационального зерна», необходимого и достаточного для построения строгих математических моделей произвольной предметной области (проблемы, зада-

чи) и их интерпретации для получения конкретных результатов.

Результаты анализа работ зарубежных ученых и синтез варианта компьютерной ОВУ изложены в работах [17, 19–21]. На его основе зафиксируем некоторую совокупность подходов, принципов и механизмов синтеза *системы категорий* как знаний об общей структуре действительности. Система категорий определена как последовательность таких ступенек познания, которые должны быть пройдены в исследовании объектов окружающего мира. Эти ступеньки познания, в свою очередь, отражают общие действия, которые совершает субъект при освоении объективного мира и которые он может применить в конкретной предметной области, исходя из существенных общих представлений о нем. Осознание того, что категории отражают всеобщие характеристики любого объекта, описывают не только общую структуру взаимодействия субъекта с объектом, но и общую структуру взаимодействия любых объектов, достигается опосредствованно, как вывод, следующий из доказательства того, что деятельность может быть успешной лишь в том случае, когда она подчиняется общим законам взаимодействия известной нам Вселенной [3]. «Система категорий выступает по отношению ко всякому новому знанию как некоторый фиксированный *категориальный каркас*, который имеется до познания данного конкретного класса явлений и определяет формы этого познания, так что всякое новое содержание, доставляемое конкретным исследователем, подводится под одну из категорий и упорядочивается по правилам категориального каркаса» [22, 23].

Результатом системного анализа, в частности, являются приведенные далее утверждения, описывающие *методологические основы исследования и систематизации категорий* [3, 17, 22–28]:

- единство онтологического и гносеологического подходов;
- гносеологический принцип субординации категорий есть движение от абстрактного к конкретному;

- рассмотрение категорий как ступеней познания;
- необходимость строгих определений категорий;
- необходимость общего понятийного каркаса как основы интерпретации частных научных теорий;
- в системе категорий формируются предельно общие принципы, которые должны быть положены в основу систематизации знаний в целом;
- категориальные отношения субъекта к объекту есть проявление категориальных отношений между любыми объектами;
- возможно построение различных систем категорий (допущение множественности систем), отражающих категориальные каркасы структур знаний, образующихся при решении научно-научных задач; эти различные системы категорий должны быть субординированы через единый базовый *неопределяемый минимум понятий*, составляющий базовую систему, дающую общий язык для построения других систем (онтологию категорий верхнего уровня);
- только с помощью *базовой системы категорий* можно решить следующие общенаучные и внутрифилософские задачи:
 - выделить семантические множители для всеобщих понятий единого языка науки;
 - последовательно и строго определить всеобщие категории, исходя из неопределяемого минимума;
 - дать основу для построения онтологической атрибутивной модели материи на уровне, отраженном современной наукой;
 - основу для системы логических операций по их направленности на познание всеобщих сторон любого объекта;
 - основу для системного построения общей методологии;
 - дать точный категориальный язык для более детальных категориальных исследований;
- определения категорий в ОВУ носят функциональный характер, т.е. они отличаются одна от другой по специфической функции в теории познания. Общий характер движения мысли в процессе выведения категорий может быть

определен как восхождение от абстрактного к конкретному; каждая последующая категория более конкретно описывает предмет, а их система в целом суть самое конкретное описание самого абстрактного уровня знаний о любом объекте нашего опыта;

- всеобщность категорий не означает их тавтологичности.

Объединенные из различных источников принципы и механизмы построения системы категорий позволяют синтезировать *методологические основы проектирования компьютерной ОВУ*. При этом следует принять, что:

- *неопределяемый минимум* базовой системы категорий имеет минимальное количество логических аксиом $((A_i), i = \overline{1,8})$ всеобщего характера, например таких, как для категорий *Абстрактное* и *Материальное* [17, 19]:

A1) не существует никакой абстрактной сущности x , размещенной в пространстве

$$\neg(\exists x: \text{Абстрактное})(\exists y: \text{Место}) \text{loc}(x, y),$$

где $\text{loc}(x, y)$ – соответствующий предикат местонахождения x в y ;

A2) не существует никакой абстрактной сущности x , которая происходит в момент времени t

$$\neg(\exists x: \text{Абстрактное})(\exists t: \text{Время}) p_{\text{time}}(x, t),$$

где $p_{\text{time}}(x, t)$ – соответствующий предикат происхождения x в t ;

A3) что-нибудь материальное x существует в некотором месте y

$$(\forall x: \text{Материальное})(\exists y: \text{Место}) \text{loc}(x, y),$$

где $\text{loc}(x, y)$ – соответствующий предикат местонахождения x в y ;

A4) что-нибудь материальное x происходит в момент времени t

$$(\forall x: \text{Материальное})(\exists t: \text{Время}) p_{\text{time}}(x, t),$$

где $p_{\text{time}}(x, t)$ – соответствующий предикат происхождения x в t ;

- наивысшим уровнем категоризации, не имеющим никакой дифференциации, является некоторая категория (названная «Универсум»), которой удовлетворяют следующие аксиомы:

A5) существует такой x , что

$$(\exists x) \text{Универсум}(x);$$

А6) всякий x есть элементом Универсума:

$$(\forall x) \text{Универсум}(x);$$

А7) всякая категория k есть подкатегория Универсума:

$$(\forall k: \text{Категория}) k \leq \text{Универсум};$$

А8) все подкатегории определяются дополнительными отличиями в Универсуме, чтобы показать, как они отличаются от Универсума и одна от другой;

• категории на «ступеньках познания» структурируются в соответствии с категориальными отношениями частичного порядка при удовлетворении условий простого и множественного наследования. Другими словами, в качестве «ступенек познания» выступает ориентированный граф с (возможными) перекрестными ребрами.

Следует отметить, что рабочей группой по разработке проекта стандарта ОВУ (*SUO WG*) был принят аналогичный подход, в результате которого была разработана онтология *SUMO* [29]. Языком представления знаний в *SUMO* был выбран *KIF*-язык, который после доработок получил название *SUO-KIF*. *SUMO* включает в себя две части – онтологию верхнего уровня (переработанные онтологии Дж. Совы и Рассела–Норвига) и онтологию среднего уровня (реструктурированные остальные онтологии). После завершения создания каждой части в отдельности они были сведены в единую концептуальную структуру.

Онтолого-управляемая архитектура

В многочисленных зарубежных научных публикациях под онтолого-управляемой архитектурой (*Ontology-driven architecture*) понимают такую архитектуру ЗОИС, в которой база знаний ПдО представлена формальной онтологией. При этом последняя описана на некотором языке описания онтологий (список таких языков определен консорциумом *W3C* (*OWL*, *RDFS* и др.)). ЗОИС с онтолого-управляемой архитектурой классифицируют [4] по функциональному признаку на онтолого-знающие информационные системы (ОЗНИС) и онтолого-управляемые информационные системы (ОУИС). ОЗНИС характеризуется «пассивной» ролью онтологии, т.е. выполнением ею роли только ин-

формационно-справочного компонента. Любая другая подсистема только «знает» о существовании (возможно отдаленно) онтологии и может использовать ее для любой специфической цели приложения. В ОУИС онтология выполняет «активную» роль, кооперируется с другими подсистемами в текущем времени в соответствии с «высшей» общей целью ЗОИС. Важное следствие использования онтологии в текущем времени – способность коммуникации между программными агентами. Последние поддерживают связь друг с другом через сообщения, содержащие выражения, сформулированные в терминах онтологии (*онтолого-управляемая коммуникация*). Для того, чтобы программному агенту «понять» значения этих выражений, ему необходим доступ к онтологии [8].

Использование онтологии в качестве компонента интерфейса пользователя не столь очевидно, однако, очень важно, так как онтология несет семантическую информацию относительно ограничений, наложенных на классы и отношения. Пользователь может ознакомиться с онтологией, чтобы лучше понять функциональные возможности используемых ЗОИС, и сделать корректные запросы к базе данных (знаний).

Использование онтологии в качестве компонента прикладной программы также является важной частью многих ЗОИС. Эти компоненты обычно содержат фрагменты знаний предметной области, которые, по различным причинам, явно не сохранены в базе данных. Некоторые части этого знания закодированы в статической части программы в форме типа или объявления класса, другие части (подобно, например, бизнес-правилам) неявно сохранены в процедурной части программы. Возможно, в принципе, генерировать статическую часть программы со справкой онтологии.

Кроме того, онтологии, интегрированные с лингвистическими ресурсами, могут быть использованы для поддержки и развития объектно-ориентированного программного обеспечения. Можно также представить все знания предметной области неявно закодированными в прикладной программе, преобразуя программу в знание-основанную систему. В этом случае ба-

за знаний содержит ядро, представленное в виде онтологии.

Интеграция междисциплинарных знаний и развивающиеся ЗОИС

Как следует из рассмотренного анализа взаимодействия общих научных дисциплин, именно теория познания определяет основные принципы и направления исследовательского процесса в системе Λ . Следует также отметить радикальное изменение и самой системы научного познания. Размываются четкие границы между практической и познавательной деятельностью. В системе научного знания проходят интенсивные процессы дифференциации и интеграции знания, развиваются комплексные междисциплинарные и трансдисциплинарные исследования, новые способы и методы познания, методологические установки, появляются новые элементы картины мира, выделяются новые, более сложные типы объектов познания, характеризующиеся историзмом, универсальностью, сложностью организации, которые раньше не поддавались теоретическому (математическому) моделированию. Одно из таких новых направлений в современном естествознании представлено синергетикой [30].

Синергетика (от греч. $\sigma\upsilon\nu$ – «совместно» и $\epsilon\rho\omega\varsigma$ – «действующий») – междисциплинарное направление научных исследований, задача которого – изучение природных явлений и процессов на основе принципов самоорганизации систем (состоящих из *подсистем*) [31, 32]. Синергетика изначально позиционировалась как междисциплинарный подход, как принципы, управляющие процессами самоорганизации, одни и те же безотносительно к природе систем. Интерес представляют конкретные научно-практические задачи, в их конкретной постановке – разработка развивающихся знание-ориентированных систем, открытых баз знаний базовых доменов прикладных областей, их сопровождение и пополнение, разработка эффективных механизмов извлечения знаний, их обнаружения и формального представления.

Следующий шаг развития междисциплинарных научных исследований и теории баз знаний в частности – теоретически обоснованное объ-

единение (или *системная интеграция*) уже разработанных как общедоступных онтологий, так и коммерческих баз знаний для разнообразных прикладных задач, проблем, целых предметных областей и трансдисциплинарных знаний общего характера. Сущность системной интеграции сформулирована следующим образом [33]: устойчивые знания совокупности научных дисциплин можно представить в форме интегрированной иерархической сети научных теорий (разного уровня развития, содержательности и охвата действительности), составляющие которых, возможно, связаны посредством общих объектов действительности. Там же говорится о цели междисциплинарных исследований – приближение к построению общенаучной картины мира – а также о системной интеграции знаний (онтологий) как одной из важных задач в достижении указанной цели и последующем построении *Единой сети знаний* [33].

В общем виде процесс, обеспечивающий системную интеграцию множества онтологий, можно представить следующим образом:

$$\overset{\cup}{O} = \bigcup_i O_i, \quad i = \overline{1, N},$$

где \bigcup_i – знак концептуального объединения. Смысл этого знака состоит в системной интеграции исходных онтологических графов с учетом областей определений O_i ($i = \overline{1, N}$) и их взаимосвязи (взаимодействия) [21]. При этом возникает вопрос об определении некоторых общих характеристик онтографа ПДО. В частности, в [21] для этой цели используется параметр оценки сложности онтографа

$$W = \sum_i \sum_h \sum_l O_i \cdot S_{h,l},$$

где O_i – онтограф i -й предметной области, $i = \overline{1, N}$; $S_{h,l}$ – степень вершины, равная числу исходящих из нее ребер, $h = \overline{1, H}$ – количество уровней онтографа, $l = \overline{1, L_h}$ – номер вершины на соответствующем (h -м) уровне онтографа.

Обобщенная архитектура знание-ориентированной системы, эффективно реализующей процедуры обработки предметных знаний, раз-

вита БЗ и системной интеграции, представлена на рис. 4.

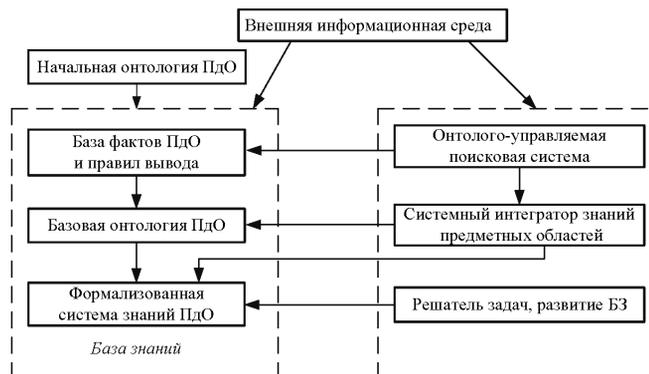


Рис. 4. Архитектура развивающихся знание-ориентированных систем

Реализация технологий представления и обработки знаний и процесса их системной интеграции предполагает учет различных формально-методологических требований, критериев и оценок. Приведем основные из них.

- Сущность методов онтологического исследования как раздела системного анализа предполагает:

- построение статических и динамических моделей;

- исследование структуры ПдО и взаимосвязи компонент знаний;

- исследование поведения системы, описывающей заданную ПдО – разработка алгоритмов процедур композиции знаний и диаграмм состояний.

- Влияние предметной области на вид онтологии.

- Правила формирования достоверных утверждений и выводов, описывающих систему и ее поведение.

- Категориальный уровень онтологии ПдО и соответствующая ему система отношений.

- Характеристики онтографа (средний коэффициент ветвления i -й вершины, количество уровней ветвления, количество типов отношений, в том числе по подтипам).

Познание идет от изучения вещей к изучению процессов (в том числе взаимодействия компонент, включая уровень «субъект–объект»), от устойчиво функционирующей системы – к системе *изменяющейся и развивающейся*, а ре-

зультат, следствие процесса развития – возникновение *нового качества* [3].

Заключение. Полученные результаты позволят повысить эффективность процедур и соответствующих алгоритмов обработки ЕЯ-информации и соответствующих инструментальных средств.

1. Палагин А.В., Крывый С.Л., Петренко Н.Г. Знание-ориентированные информационные системы с обработкой естественно-языковых объектов: основы методологии и архитектурно-структурная организация // УСиМ. – 2009. – № 3. – С. 42–55.
2. Палагин А.В., Петренко Н.Г. Системно-онтологический анализ предметной области // Там же. – № 4. – С. 3–14.
3. Сагатовский В.Н. Основы систематизации всеобщих категорий. – Томск: Изд-во Томск. ун-та, 1973. – 432 с.
4. Guarino N. Formal Ontology and Information Systems // Proc. of FOIS'98. – Trento, Italy, 6–8 June, 1998. – Amsterdam: IOS Press. – P. 3–15.
5. Gruber T.R. A translation approach to portable ontology specifications // Knowledge Acquisition. – 1993. – 5 (2). – P. 199–220.
6. Gruber T.R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // Intern. J. of Human and Comp. Studies. – 1995. – 43(5/6). – P. 907–928.
7. Niles I., Pease A. Towards a Standard Upper Ontology // Proc. of the 2nd Intern. Conf. on Formal Ontology and Inform. Syst. (FOIS-2001). Ed by Welty C., Smith B. Ogunquit, Maine. – Oct. 17–19, 2001. – 8 p. – <http://home.earthlink.net/~adampease/professional/FOIS.pdf>
8. Палагин А.В., Яковлев Ю.С. Системная интеграция средств компьютерной техники. – Винница: «УНІВЕРСУМ-Вінниця», 2005. – 680 с.
9. Башмаков А.И., Башмаков И.А. Интеллектуальные информационные технологии: Учебное пособие. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2005. – 304 с.
10. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. – 384 с.
11. Гаврилова Т.А. – Онтологический инжиниринг. – /Available at http://www.kmtec.ru/publications/library/authors/ontolog_engineering.shtml
12. Методы и средства автоматизированного проектирования прикладной онтологии / Б.В. Добров, Н.В. Лукашевич, О.А. Невзорова и др. // Изв. РАН. – Теория и системы управления. – 2004. – № 2. – С. 58–68.
13. Логические методы и формы научного познания / В.Т. Павлов, К.Ф. Руденко, И.С. Семенов и др. – К.: Вища шк., 1984. – 208 с.
14. Андон Ф.И., Яшунин Л.Е., Резниченко В.И. Логические модели интеллектуальных информационных систем. – К.: Наук. думка, 1999. – 397 с.

15. Палагін О.В., Петренко М.Г. Архітектурно-онтологічні принципи розбудови інтелектуальних інформаційних систем // Матем. маш. і сист. – 2006. – № 4. – С. 15–20.
16. Палагін А.В. Организация и функции «языковой» картины мира в смысловой интерпретации ЕЯ-сообщений // Inform. Theories and Application. – 2000. – 7, № 4. – С. 155–163.
17. Палагін О.В., Петренко М.Г. Модель категоріального рівня мовно-онтологічної картини світу // Матем. маш. і сист. – 2006. – № 3. – С. 91–104.
18. Sowa John F. Knowledge Representation: Logical, Philosophical, and Computational Foundations // Brooks Cole Publ. Co., Pacific Grove, CA, 2000. – 594 p.
19. Палагін О.В., Петренко М.Г. Розбудова абстрактної моделі мовно-онтологічної інформаційної системи // Матем. маш. і сист. – 2007. – № 1. – С. 42–50.
20. Палагін О.В., Петренко М.Г., Михайлик А.В. Розвиток та порівняльні характеристики логіко-онтологічних формальних теорій // Там же. – № 2. – С. 3–18.
21. Палагін А.В., Петренко Н.Г. К вопросу системно-онтологической интеграции знаний предметной области // Там же. – № 3, 4. – С. 63–75.
22. Рубашкин В.Ш. Проблема интерпретации в физической теории // Логика и методология науки. – М.: 1967. – С. 35–49.
23. Соколовская Ж.П. «Картина мира» в значениях слов. – Симферополь: Таврия. – 1993. – 197 с.
24. Кузьмин Е.С. Система онтологических категорий. – Иркутск: 1958. – 183 с.
25. Свидерский В.И. Некоторые вопросы диалектики изменения и развития. – М.: Мысль, 1965. – 228 с.
26. Библер В.С. О системе категорий диалектической логики. – Сталинабад: 1958. – 157 с.
27. Зиновьев А.А. Основы логической теории научных знаний. – Ин-т филос. АН СССР.– М.: Наука, 1967. – 202 с.
28. Апресян Ю.Д. Идеи и методы современной структурной лингвистики. (Крат. очерк). – М.: Просвещение, 1966. – 302 с.
29. SUO The IEEE Standard Upper Ontology 2001. – <http://suo.ieee.org>
30. Чернавский Д.С. Синергетика и информация: Динамическая теория информации. – М.: УРСС, 2004. – 288 с.
31. http://slovari.yandex.ru/dict/phil_dict/article/filo/filo-707.htm.
32. Князева Е.Н., Курдюмов С.П. Основания синергетики. Синергетическое мировидение. – М.: КомКнига, 2005. – 240 с.
33. Палагін О.В., Кургаєв О.П. Міждисциплінарні наукові дослідження: оптимізація системно-інформаційної підтримки // Вісн. НАН України. – 2009. – № 3. – С. 14–25.

Поступила 29.04.2010

Тел. для справок: (044) 526-3348 (Київ)

© А.В. Палагін, С.Л. Кривий, Н.Г. Петренко, В.Ю. Величко, 2010

Внимание !

**Оформление подписки для желающих
опубликовать статьи в нашем журнале обязательно.
В розничную продажу журнал не поступает.
Подписной индекс 71008**