

УДК 004.318

А.В. Палагин, Н.Г. Петренко

## Системно-онтологический анализ предметной области

Предложен системно-онтологический подход к построению базы знаний, функционирующей в составе онтологоуправляемой информационной системы. Компонентами этой базы знаний являются онтология объектов, онтология процессов и онтология задач, для которых разработаны схемы моделей и общая методика проектирования.

A system-ontological approach to construction of the knowledge base functioning in a structure of the ontology-driven information system is suggested. The components of this knowledge base are the ontology of objects, the ontology of processes and the ontology of tasks for which the models of schemes and a general technique of designing are developed.

Запропоновано системно-онтологічний підхід до побудови бази знань, яка функціонує у складі онтологокерованої інформаційної системи. Компонентами цієї бази знань є онтологія об'єктів, онтологія процесів та онтологія задач, для яких розроблено схеми моделей і загальна методика проектування.

Согласно [1] под *gebahf* понимается вид исследования, при котором реальный или мыслимый объект расчленяется на составляющие части (элементы) и исследуются эти элементы и связи между ними. Анализ предметной области (ПдО) представляет особый вид научной деятельности, в результате которой строится интерпретационная модель предметных знаний (в широком смысле). В процессе анализа знания делятся на инвариантные и прагматические, концептуальные составляющие которых представляют онтологические знания ПдО.

Некоторые идеи по разработке методологии проектирования онтологии ПдО берут свое начало в литературе по объектно-ориентированному подходу (ООП), возникшему как технология программирования больших программных продуктов [2]. Однако разработка онтологий как иерархической структуры понятий (концептов) отличается от проектирования объектов как классов и отношений в объектно-ориентированном программировании. Последний сосредотачивается главным образом на методах описания классов – программист принимает проектные решения, основанные на *hijuo* свойствах класса, тогда как разработчик онтологии принимает эти решения, основываясь на *kljmdlmjguo* свойствах класса. В

результате структура понятий и отношения между понятиями в онтологии отличается от структуры классов объектов подобной ПдО в объектно-ориентированной программе [3]. Кроме того, при разработке онтологии содержимое понятия эксплицируется всегда, в то время как в объектно-ориентированном программировании зачастую применяется метод инкапсуляции как способ ограничения доступа к внутреннему содержимому объекта.

Системный подход к познанию ориентирует аналитика на рассмотрение любой ПдО с позиций закономерностей системного целого и взаимодействия составляющих его частей. Системность знаний исходит из многоуровневой иерархической организации любой сущности, т.е. все объекты, процессы и явления можно рассматривать как множество более мелких подмножеств (признаков, деталей) и, наоборот, любые объекты можно (и нужно) рассматривать как элементы более высоких классов обобщения.

Начало 90-х годов прошлого столетия считается зарождением парадигмы компьютерных онтологий, которая была сформулирована как попытка сгладить (и по возможности устранить) проявляющиеся разного рода противоречия при функционировании и внедрении интеллектуальных систем с использованием баз знаний

ПдО. Ярким представителем таких систем на то время были экспертные системы (ЭС). Они успешно и эффективно функционировали в пределах одного коллектива и на уровне коммерческих образцов. У разработчиков постоянно возникал вопрос: «Как обеспечить продвижение этих образцов к конечному пользователю?». Были предложены разработки оболочек – «пустых» ЭС и ряд других новшеств. Но они не имели решающего значения. Напрашивался вывод, что для эффективного функционирования ЭС на самом важном этапе «жизненного цикла» – функционировании у конечного пользователя при решении реальных задач – необходимо к каждой ЭС «приложить» эксперта в соответствующей ПдО. Необходимость присутствия эксперта объяснялась, в частности, быстротечностью изменения знаний во многих предметных областях и соответственно необходимостью обновления базы знаний ЭС в «реальном времени». В тех же ПдО, где знания относительно инвариантны, ЭС продолжали эффективно функционировать.

Изложенное выше и ряд других факторов «подтолкнули» ученых к разработке парадигмы  $\text{dhfivxljguo} \quad \text{hglheh\#c}$ , основные принципы которой сформулированы в [4].

1.  $\text{hohqbhklv}$ ,  $\text{ykgghklv}$  (*Clarity*). Термины (и понятия) онтологии должны отражать реальную действительность. Их символьные обозначения (знаки) должны формироваться на основе общепринятых правил в семиотике и выражать общепринятые смыслы реальных объектов. Эти смыслы, в свою очередь, извлекаются из общепринятых определений терминов (понятий), зафиксированных в толковых словарях и различных глоссариях ПдО. Суждения, входящие в определения, формализуются на основе формального общепринятого аппарата в виде тождественно истинных логических аксиом.

2.  $\text{H[hkghgghklv}$ ,  $\text{kyaghklv}$  (*Coherency*). Формирование начального набора понятий онтологии и их добавление должно быть обоснованным, определяемым, в первую очередь, требованиями предполагаемой совокупности ре-

шаемых задач. Логические аксиомы начального набора понятий должны быть непротиворечивыми. Для этого необходим предусмотренный механизм логического вывода, который также проверяет на непротиворечивость добавляемые аксиомы и выводимые в онтологии утверждения.

3.  $\text{Jkfbjy\_fhklv}$  (*Extendibility*). Ядром онтологии являются первоначально введенные (спроектированные) понятия и описывающие их аксиомы. В онтологии должен быть предусмотрен механизм расширения (ограничения) совместно используемых словарей понятий без нарушения целостности системы.

4.  $\text{Fbgbfvgh} \square \quad \text{\`bygb\_ dhhjh\#by}$  (*Minimal encoding bias*). В онтологической системе (ОнС) должен быть реализован принцип совместного использования онтологий, который предполагает спецификацию онтологии на уровне представления, а не символьного кодирования, запись которой на общепринятом и платформонезависимом языке описания онтологий можно передать для использования любому программному агенту.

5.  $\text{Fbgbfvgu} \square \quad \text{hglheh\#qkdb\_ h\#a\#ev} - \text{klZ}$  (*Minimal ontological commitment*). Этот принцип перекликается с принципами обоснованности и расширяемости/ограничения. Необходимо, чтобы множество понятий онтологии отображало концептуальную структуру ПдО, относительно стабильную на протяжении «жизненного цикла» ОнС. А последняя предоставляла бы возможность расширения или специализации отдельных ветвей онтологического графа (онтографа, ОГ). Отделение концептуальных знаний от знаний, выраженных фактами, является стратегией построения ОнС, а точнее – онтологических баз знаний.

$\text{Ihklghd} \square \quad \blacksquare$

Известны общие совокупности методов, принципов, процедур и атрибутов системного анализа как научного познания в любой ПдО [5, 6 и др.]. При конкретизации ПдО конкретизируются и средства системного анализа. Парадигма компьютерных онтологий, развиваемая во взаимодействии со средствами и методами системного анализа, положила начало развитию но-

вой ветви средств и методов системного анализа ПдО – системно-онтологическому анализу (или подходу).

Решение задачи (СОП) есть разработка онтологических средств поддержки решения прикладных задач –  $ih\bar{e}b\bar{n}m\bar{g}d\bar{p}b\bar{h}g\bar{e}v\bar{g}h\bar{c} \quad hgl\bar{h}e\bar{h}\bar{f}q\bar{c} \quad kd\bar{h}\bar{h} \quad i\bar{h}\bar{d}\bar{h}\bar{z}$  (СОП) есть разработка онтологических средств поддержки решения прикладных задач –  $ih\bar{e}b\bar{n}m\bar{g}d\bar{p}b\bar{h}g\bar{e}v\bar{g}h\bar{c} \quad hgl\bar{h}e\bar{h}\bar{f}q\bar{c} \quad kd\bar{h}\bar{c} \quad kb\bar{k}l\bar{f}u$  <sup>1</sup>. Такая система (точнее, ее концептуальная часть) описывается кортежем (1) и представляет онтологию ПдО, состоящую из онтологии объектов и онтологии процессов, и онтологию задач [7–8].

$$ОнС = \langle O^{ПдО} (O^O, O^П), O^3 \rangle. \quad (1)$$

На рис. 1 представлена схема компонентов проблемного пространства (ПрП) и соответствующие ему онтологии. ПрП – это модель всех таких аспектов или компонент ПдО, с которыми связаны (опосредованно или непосредственно) знания, необходимые при решении различных задач в этой ПдО. Всякое ПрП состоит из двух блоков: инвариантной (относительно неизменной) части и множества изменяемых частей, соответствующих отдельным задачам. В составе инвариантной части, например в методологии *SMEE (Structured Methodology for Elicitation of Expertise)*, выделяют семь типов компонент: объекты, инструменты, операторы, операции, конечные продукты, побочные продукты и ограничения [9]. Эти типы компонент – суть понятия, которые хорошо группируются в онтологии объектов и процессов, представленных на рис. 1, где приняты следующие обозначения:

- $O^O$  – онтология множества объектов (понятий, концептов) ПдО, рассматриваемая как иерархическая структура классов, подклассов и элементов классов;
- $O^П$  – онтология множества процессов ПдО, рассматриваемая как иерархическая структура процессов, подпроцессов, действий и операций;

- $O^3$  – онтология совокупности задач, которые могут быть поставлены и решены в ПдО; рассматривается как иерархическая структура задач, подзадач, процедур и операторов.

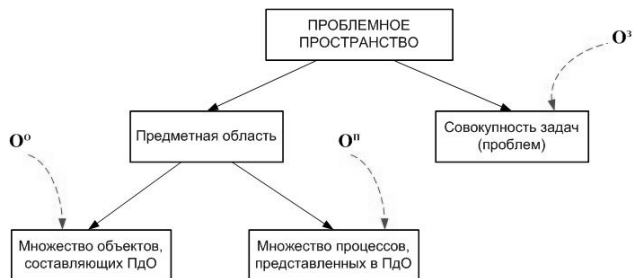


Рис. 1. Схема компонентов проблемной области

Итак, решение задачи является разработкой формализованной методики проектирования онтологических знаний ПдО, состоящих из онтологии объектов, онтологии процессов и онтологии задач. Такая методика – существенная составляющая при построении онтологизированных систем знаний предметных областей в рамках программы «Информационные технологии – Междисциплинарные научные исследования» [10].

### Исходные данные

Известна методология структурного анализа и проектирования (*SADT*) сложных систем в произвольной предметной области [11]. Эта методология породила семейство методик (и соответствующих стандартов) *IDEF (Integrated DEFinition)*, ориентированных на разработку моделей ПдО и акцентирующих внимание на каком-то конкретном аспекте проектирования [12–13]. В частности:

- методика *IDEF0* рекомендована для содержательного анализа и функционального проектирования сложных систем управления, в том числе и программного обеспечения; однако в последнем случае трудно определить ее преимущества в сравнении с методологией ООП. Описание объектов и процессов в методике выполняется в виде иерархической совокупности диаграмм с лаконичным описанием функций; блоки на диаграммах выражают функции, поэтому их названия – глаголы или отглагольные существительные;

<sup>1</sup> В данной статье не рассматриваются такие компоненты онтологической системы, как машина вывода и метаонтология.

• методика *IDEF1* предназначена для информационного моделирования, основана на концепции «сущность–связь»; обычно отправным пунктом для разработки информационной модели является *IDEF0*-модель;

• методика *IDEF3* описывает поведенческие аспекты конкретных приложений, рассматривает последовательность выполнения и причинно-следственные связи между ситуациями и событиями для структурного представления знаний о ПдО; если *IDEF0* связана с функциональными аспектами и отвечает на вопрос «Что делает система?», то в *IDEF3* детализируются *IDEF0*-функции. Эта модель отвечает на вопрос «Как система это делает?»;

• методика *IDEF5* предназначена для онтологического анализа ПдО, анализа основных терминов и понятий, используемых для описания объектов и процессов, границ использования, а также взаимосвязей между ними. Служит для эффективного исследования и документирования: словаря терминов, используемых при описании характеристик объектов и процессов, имеющих отношение к рассматриваемой ПдО, точных и однозначных определений всех терминов и классификации логических взаимосвязей между ними.

Для каждой из методик семейства *IDEF* разработаны этапы и стадии построения модели ПдО, языки и диаграммы представления результатов. Анализ получаемых в результате применения методик *IDEF* моделей ПдО и их описаний с точки зрения формализации и компьютерной обработки показал, что, по сути, они (функциональные модели и описания) представляют собой содержательное техническое задание на проектирование ОнС, включающей онтологию объектов (сущностей), онтологию процессов и онтологию задач. Описание собственно ОнС имеет несколько иные цели. Оно ориентировано на компьютерное представление с помощью стандартизованных языков предметных знаний с целью широкого использования сообществом пользователей.

Итак:

• методики *IDEF* и системно-онтологический подход используют одни и те же множест-

ва сущностей ПрП, акцентируя внимание на различных совокупностях характеристик и атрибутов;

• конечные цели этих двух подходов отличаются: для *IDEF* – функционально-блочная модель, а для СОП – онтолого-содержательная модель.

Изложенный анализ позволяет утверждать, что методики *IDEF* и системно-онтологический подход решают разные задачи, но со значительным «перекрытием» некоторых функций. Это проявляется для пары «онтология процессов ↔ поведенческая модель ПдО». В действительности в ООП (как и в онтологии задач) основными понятийными единицами являются классы объектов (задач) и методы их описания (решения).

Очевидно, этапы проектирования онтологических систем исходят из принятых в методологии *SADT* фаз проектирования сложных систем:

- *g\_e b a* – определение того, что система будет делать;
- *i j h d l b j h g b* – определение подсистем и их взаимодействие;
- *j e b a p b y* – разработка каждой подсистемы;
- *h t b g g b* – соединение подсистем в единое целое;
- *l k l b j h g b* – проверка работы системы;
- *m k l g h d* – введение системы в действие;
- *n m g d p b h g b j h g b* – использование системы.

### **Предварительный анализ предметной области**

Во все методологии включен этап предварительного анализа ПдО или составления содержательного ТЗ на проектирование [11–18 и др.]. Этот этап (как и процесс проектирования базы знаний ПдО в целом) носит сложный аналитический характер и заключается в многократном абстрагировании, в результате которого из всего многообразия сторон и свойств сущностей предметной области выделяются наиболее существенные, релевантные конкретным задачам. Знание ПдО, понимание сути происходящих в ней процессов, законов, правил и ог-

раничений, управляющих ее развитием, есть необходимое условие успешного решения задач. Более того, наличие таких знаний – необходимое условие постановки, формулировки этих задач, без чего невозможно решение [19]. Онтологические системы призваны сделать знания коллективным достоянием, дать мощный инструмент для фиксации, приобретения и обработки знаний, проверки их на непротиворечивость, полноту и пр. Кроме того, составляется систематизированное представление знаний о ПдО, выявляются источники формирования элементов множеств и процедур, задач, выполняемых в анализируемой ПдО. Составляется и документируется словарь терминов ПдО.

Этап предварительного анализа ПдО включает в себя:

- обоснованный выбор точного (и достаточного) фрагмента ПрП, относительно которого будут ставиться и решаться задачи пользователя;

- выбор методов и процедур системно-онтологического анализа, которыми, в частности, могут быть абстрагирование и конкретизация, композиция и декомпозиция, структурирование, кластеризация и классификация, тестирование и верификация;

- составление детального словаря терминов и его разбиение на подмножества терминов–объектов, терминов–процессов и терминов, имеющих задачи и методы.

Если проблемное пространство представляет собой сложную систему, то следует рассмотреть вопрос о предварительном этапе проектирования на основе методик *IDEF*, которые дополняют описанные шаги проектирования. Тем более, что известные инструментальные средства анализа ПдО и построения баз знаний (например, *KADS*-системы, российская *SIMER+MIR*) по ряду причин не могут быть использованы.

Как правило, методика сводится к алгоритму итеративного характера. Для процесса разработки необходимо предусмотреть ряд «контрольных точек» проверки полученных результатов на соответствие выбранным критериям.

Указанные критерии должны соотноситься с заданными критериями на проектирование базы знаний ПдО, так как ее создание – цель разработчиков. Оптимальный результат, как правило, зависит от степени проработки предполагаемых приложений и вариантов использования онтологии.

#### Онтология объектов □

Под  $h\bar{g}l\bar{h}e\bar{h}j\bar{b}\bar{c} \quad h\bar{f}d\bar{l}h\backslash$  предметной области понимается кортеж четырех множеств:

$$O^0 = \langle X, R, F, A(D, R_s) \rangle, \quad (2)$$

где:

- $X = \{ x_1, x_2, \dots, x_i, \dots, x_n \}$ ,  $i = \overline{1, n}$ ,  $n = Card X$  – конечное множество концептов (понятий) заданной ПдО;

- $R = \{ r_1, r_2, \dots, r_k, \dots, r_m \}$ ,  $R: x_1 \times x_2 \times \dots \times x_n$ ,  $k = \overline{1, m}$ ,  $m = Card R$ , – конечное множество семантически значимых отношений между концептами ПдО. Они определяют тип взаимосвязи между понятиями. В общем случае, отношения делят на  $h[s\bar{a}g\bar{t}b\bar{f}u]$  (из которых выделяют, как правило, отношения частного порядка) и  $dh\bar{g}d\bar{j}\bar{l}g\bar{u}] \quad h\bar{l}g\bar{h}r\bar{g}b\bar{y}$  заданной ПдО;

- $F: X \times R$  – конечное множество функций интерпретации, заданных на концептах и/или отношениях;

- $A$  – конечное множество аксиом, используемых для записи всегда истинных высказываний (определений и ограничений).

Компьютерная онтология – выражение (формальное) концептуальных знаний о предметной области и по своей значимости сопоставима с базой знаний интеллектуальной информационной системы, а ее построение есть специфическая форма человеческого творчества. Творческий процесс можно представить совокупностью операций-процедур с суждениями, утверждениями, понятиями и отношениями между ними. А последние являются фундаментом, основой для построения составной части научной теории – онтологической базы знаний в заданной предметной области. При этом такие знания описываются в декларативной форме.

Онтология определяет общеупотребительные, семантически значимые  $ihgy|bcgu$   $\mathfrak{h}$  -  $gbpu agzbc$ , которыми оперируют исследователи и разработчики знаниеориентированных информационных систем. Она отделяет  $kl$  -  $lbqkdb$  и  $bgfbq|kdb$  компоненты знаний ПдО от операциональных знаний. В отличие от знаний, закодированных в алгоритмах, онтология обеспечивает их унифицированное и многократное использование разными исследовательскими группами, на разных компьютерных платформах при решении разных задач.

$lhkljh|gb$   $\mathfrak{h}$   $\mathfrak{h}$

Напомним некоторые определения, непосредственно относящиеся к построению множеств концептуальной модели ПдО или ее онтологии [20, 21].

**Понятие** есть целостная совокупность суждений, в которых что-либо утверждается об отличительных признаках исследуемой сущности, ядром которой являются суждения (или утверждения) о наиболее общих и в то же время существенных признаках этой сущности.

Каждое понятие характеризуется объемом и содержанием – двумя взаимосвязанными сторонами понятия. **Объем** – класс обобщенных в понятии предметов, **содержание** – совокупность (обычно существенных) признаков, по которым произведено обобщение и выделение предметов в данном понятии. Объем понятия является определяющим при формировании иерархической структуры соответствующего онтографа, а содержание – при аксиоматизации его (ОГ) вершин.

Все понятия (или концепты) делятся на ряд классов по семантической зависимости:

- от образа вида или рода предметов – на  $bhu$  и  $jihu$  понятия;
- от образа части или целого предметов – на  $ihgy|lby$  -  $qzlb$  и  $ihgy|lby$  -  $peu$  ;
- от количества отображаемых предметов – на  $bgbcgu$  и  $h|sb$  понятия;
- от отображения предмета или свойства, абстрагированного от предмета, – на  $dhgdj|lgu$  и  $[kljdlgu$  понятия.

Онтология ПдО – это концептуальная модель реального мира и ее понятия должны отражать эту реальность.

$lhkljh|gb$   $\mathfrak{h}$   $\mathfrak{h}$   $X$  считается наиболее важным моментом при разработке онтологии ПдО. Оно должно быть обязательно непустым. Соотношения между  $Card X$ ,  $Card R$   $Card F$  характеризуют онтологию по функциональному признаку.

Для хорошо проработанных предметных областей за основу множества элементов  $\{x_i\}$  может быть принято содержимое подходящих толковых словарей. В противном случае следует составить полный список терминов, в котором указать:

- чем является каждый термин – понятием-классом предметов или конкретным понятием;
- для каждого термина возможные существенные отношения с другими терминами из списка;
- описать возможные существенные свойства понятий.

Пересечение объемов и содержаний понятий в таком предварительном списке не существенно)

Известно, что в любой предметной области существуют термины-синонимы. Для них в онтологии отводится только одно понятие, в аксиомах которого может быть указан синонимический ряд терминов, т.е. синонимы одного и того же понятия не представляют различные классы.

Далее следует уточнить и определить окончательный список классов-понятий, имена которых будут входить в разрабатываемую онтологию и являться вершинами онтографа. Также следует принять единые  $ijbe$   $ijbkz$  -  $zby$   $bfg$  понятиям и их свойствам.

Затем, возможно, следует повторить некоторые фрагменты процесса анализа ПдО (с привязкой к составленному списку понятий), выполненные на предварительном этапе. Отметим, что в число указанных  $dhgljhevguo$   $lh$  -  $qd$  (точек входа итерации) должно быть включено завершение разработки любого компонента онтологии.

В результате должен быть получен полный список существенных для заданной ПдО (и предполагаемых приложений) понятий и их машинно-интерпретируемые формулировки.

$ihkljh\_gb$   $R$  также основано на результатах предварительного этапа анализа ПдО. По сути, требуется установить для каждого элемента  $x_i \in X$  семантическое отношение  $R_k$  с элементом  $x_j \in X : x_i R_k x_j, i, j = \overline{1, n}, i \neq j, k = \overline{1, m}$ . Другими словами, необходимо построить множество ребер, связывающих узлы  $ghlgh$   $hgh$ . В качестве узлов онтографа выступает множество понятий ПдО. Вершиной (или вершинами) онтографа является родовое понятие, которое не имеет надкласса, а самый нижний уровень представляют конкретные понятия, т.е. не имеющие видовых понятий в заданной ПдО.

На практике множество  $R$  на начальном этапе представляют некоторым обобщенным отношением  $uf$  –  $gb$ . Известно несколько подходов для разработки иерархии классов: процессы  $gbkohys\_c$ ,  $hkohtys\_c$  и  $dhf[bgb] - jhghc$  разработки. Последний наиболее часто используется разработчиками, так как он более естественный и сначала оперирует понятиями среднего уровня, к которым наиболее часто обращаются разработчики. Затем эти понятия обобщаются и ограничиваются.

В заключение данного подэтапа следует соотнести разработанные классы и их иерархии с результатами предварительного анализа ПдО. В частности, уточняются зависимости для конкретных пар  $(x_i, x_j)$ . В процессе соотнесения (и построения иерархии) следует учитывать, что [3]:

- прямые подклассы в иерархии должны располагаться на одном уровне обобщения;
- класс может быть подклассом нескольких классов, и тогда он может наследовать свойства от всех этих классов;
- если класс имеет только один прямой подкласс, то, возможно, при моделировании допущена ошибка или онтология неполная;

- если у данного класса есть более дюжины (иногда семь) подклассов, то, возможно, необходимы дополнительные промежуточные классы;

- в онтологии число классов соотносится с числом предполагаемых приложений.

Следует помнить, что не существует единственно правильной иерархии классов.

Описанное построение онтографа – специальный вид классификации понятий ПдО –  $hglheh]bq\_kdhc$   $deZkbnbdzbc$ .

$ihkljh\_gb$   $F$   $A$ . Известные методики разработки онтологии [3, 7, 14–18 и др.] предлагают различные толкования этих множеств. В зависимости от функциональной ориентации проектируемой онтологии множества  $F$  и  $A$  могут интерпретироваться по-разному:

- $A \equiv F$  – множество аксиом тождественно множеству функций интерпретаций. В этом случае устанавливаются существенные связи между разрабатываемыми компонентами онтологии и вариантами её использования. Основным назначением такой онтологии является однозначная интерпретация понятий, входящих в онтологию, сообществом пользователей;

- $A \neq F$  – множество аксиом не тождественно множеству функций интерпретаций; в аксиомах задаются: базовые функции (подмножество  $F$ ) либо дополнительные отношения (не являющиеся элементами множества  $R$ ) между понятиями, ограничения и условия, которые анализируются в машине вывода ОнС и используются в процессе решения задач;

- $A \neq F$  – множество аксиом не тождественно множеству  $F$ ; функции интерпретации рассматриваются как специальный вид отношений на множестве понятий  $F : x_1 \times x_2 \times \dots \times x_{n-1} \Rightarrow x_n$ . В этом случае устанавливаются существенные связи между разработанными компонентами онтологии и совокупностью задач предполагаемого приложения (приложений). Онтологии с таким представлением  $F$  используются в вопросно-ответных системах, в которых результатом является одно из значений двухэлементного множества  $\{ \quad , \quad \}$  или имя предиката.

В конечном счете, независимо от того, какая из этих формулировок будет принята, эффективность разработанной онтологии будет определяться конечными результатами приложений.

В определении концептуальной модели онтологии объектов множество аксиом  $A$  состоит из множества определений  $D_i^f$  и множества ограничений  $Rs_i^f$  для понятия  $X_i$ . Определения записываются в виде тождественно истинных высказываний, которые могут быть взяты, в частности, из толковых словарей ПдО. В них могут быть указаны дополнительные взаимосвязи понятий  $X_i$  с понятиями  $X_j$ . Во множестве ограничений  $Rs_i$  могут быть заданы ограничения на интерпретацию соответствующих понятий  $X_i$ .

Кроме того, из полного списка отобранных в онтологию терминов не все представляют понятия. Существуют термины (например, ролевые), соответствующие свойствам определенных классов-понятий. Такие свойства следует привязать к описанию самого общего класса, обладающего ими. А подклассы этого класса будут наследовать указанное свойство (конечно, если между ними установлено некоторое отношение частичного порядка).

$Khckl$  и  $ihgylbc$  имеют определенные значения, такие, как  $lbi$  значений,  $fhsghklv$  значений,  $ja[r]ggü$  значения (для данного класса) и другие. Например, мощность значений можно описать: с единичной мощностью, мощностью без ограничений и мощностью с некоторым допустимым интервалом.

На основе построенных множеств кортежа можно синтезировать концептуальную модель ПдО, например, с помощью известного инструментального средства *Protégé* и получить формальное описание разработанной онтологии на одном из языков описания онтологий, а также графическое представление онтографа.

### Ng[heh]by и N

Синонимами онтологий объектов и процессов являются соответственно статическая и динамическая онтология ПдО. В научно-технической литературе, когда говорят об онтологии

ПдО, подразумевается ее статическая составляющая. Именно компоненты последней наиболее разработаны как в литературе по философии, так и в конкретных описаниях ряда предметных областей. Под поведенческим описанием сущностей-процессов чаще всего понимают разработку определенных бизнес-процессов. Их результат – графические диаграммы и естественно-языковые описания процессов. Разработка же базы знаний не является прямой целью указанных методик. Поэтому методики разработки онтологии процессов практически неизвестны. Хотя следует отметить, что в некоторых известных онтологиях верхнего уровня (например, онтология Дж. Совы, *SUMO* и *Mikrokosmos*) сущность понятия «Процесс» рассмотрена детально [22–25].

На рис. 2 представлен синтезированный онтограф (на основе анализа указанных онтологий) той части схемы начального развития понятия «Процесс», которая соответствует процессам в научно-технических предметных областях. Ветви онтографа «Социальный процесс», «Материальный процесс» и аналогичные им не рассматриваются.

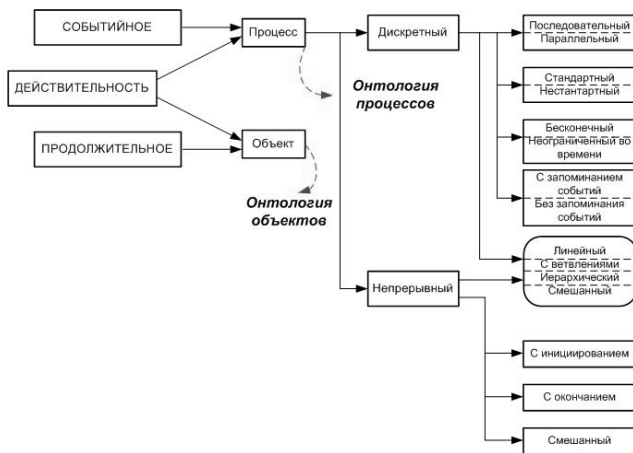


Рис. 2. Начальное развитие онтологии процессов

Категория  $Ijhp\_kk$  рассматривается как  $cklbi\_evghklv$  и  $Kh[ulbcghklv]$ , в отличие от категории  $H[i]dl$ , характеризующийся как  $cklbi\_evghklv$  и  $Ijhhebl\_evghklv$  [26]. В первую очередь  $Ijhp\_kk$  рассматривается как времязависимая категория и затем подразделяется по видам изменений, наличием



начальных и конечных точек и т.д. Далее  $[j]_k$  -  $p$  подразделяется на непрерывный и дискретный. Первый из них характеризуется наличием эксплицитных начальной и конечной точек или как без явного указания этих точек. Второй вид процесса указывает, что изменения происходят дискретными шагами, названными событиями, которые чередуются с периодами покоя, названными состояниями.

Приведенная схема не отражает всех характеристик (оснований ветвлений в онтографе) категории  $[j]_k$ , даже для той ее части, что представлена на рис. 2.

На рис. 3. представлена общая схема онтологии процесса ПдО, в которой категория  $[j]_k$  представлена онтографом с  $p$  уровнями и  $n_p$  подпроцессами (ПП) на каждом уровне. Предпоследний ( $p-1$ ) уровень представлен множествами действий, на которые разбивается каждый ПП предыдущего уровня. В свою очередь, каждое действие на последнем ( $p$ -м) уровне разбивается на последовательность операций  $O_{ij}^p$ .

Связи между подпроцессами для соседних уровней соответствуют отношениям «Целое-часть», а внутри каждого уровня – некоторой смешанной форме организации соединений. На рис. 3 показан частный случай такой организации – параллельный. Дальнейшее развитие (конкретизация) онтологии процессов возможно, когда задана конкретная предметная область и соответствующее проблемное пространство, а в более узком смысле – конкрет-

ные признаки ветвлений (условия инициирования ПП, условия окончания ПП и ограничения) в онтографе.

Объем динамических знаний  $W$  в предметной области можно оценить через характеристики (параметры) их формально-онтологических представлений. В частности, при представлении онтографом (без учета сложности функций интерпретации подпроцессов) величина  $W$  может характеризоваться числом вершин ОГ. Для ОГ, представленного на рис. 3, это число может быть выражено формулой

$$W = \sum_{k=1}^p \sum_{i=1}^n \sum_{j=1}^m \text{ПП}_{ij}^k, \quad (3)$$

где ПП – подпроцесс (вершина ОГ), Д – действие (вершина ОГ на  $(p-1)$ -м уровне), О – операция (вершина ОГ на  $p$ -м уровне);

- $k = \overline{1, p}$ ,  $p$  – количество уровней в онтографе процесса;
- $i = \overline{1, n}$ ,  $n$  – количество подпроцессов на  $i$ -м уровне;
- $j = \overline{1, m}$ ,  $m$  – количество подпроцессов в  $i$ -м подпроцессе  $(i-1)$ -го уровня.

### Нгльгьby 4 J

В текстовых описаниях (спецификациях) целевых задач выделяются набор объектов и набор процессов (методов), необходимых и достаточных для выполнения конкретных целевых задач. Можно выделить некоторое унифицированное (пополняемое) множество базовых задач (типовых фрагментов задач), на основе которых с помощью определенных логических

последовательностей конструировать более сложные задачи. Таким образом, решение прикладной задачи включает в себя, в том числе взаимодействие трех онтологий: объектов, процессов и базовых задач [7, 8].

Поле конструктора, описывающего схемы типов решаемых задач, является некоторой семантической осью

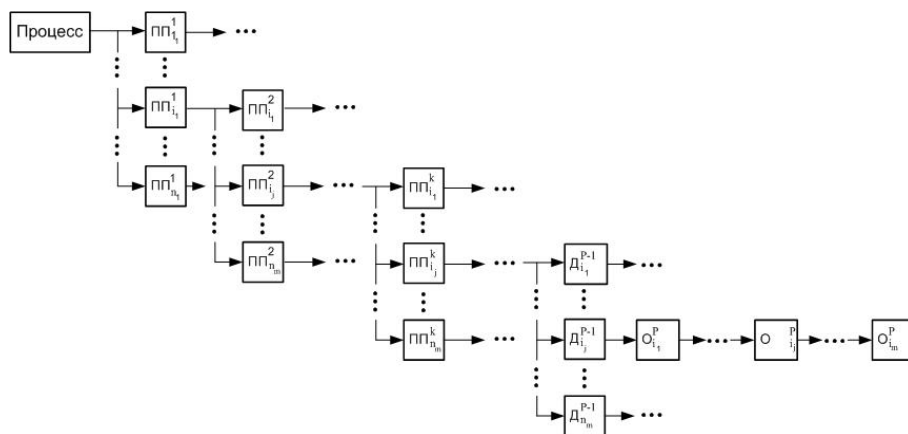


Рис. 3. Общая схема онтологии процесса

с противоположными сторонами «Выбор» и «Построение», концепты которых описывают простую и сложную схемы решения задач. Под простой схемой «Выбор» понимается такая, в которой для решения задачи онтологической системе необходимо сделать выбор из совокупности известных схем решений. Когда выбор сделан, то известны метод, процесс и его состояния, т.е. решение задачи становится тривиальным. Противоположностью простой схемы является сложная схема «Построение», когда все составляющие процесса решения задачи неизвестны. На практике, чаще всего, встречаются некоторые промежуточные схемы решения задач. Например, схемы, в которых начальные и целевые состояния, совокупности методов задаются в явном виде. Общей стратегией решения сложных (непростых) задач является многоуровневая декомпозиция исходной задачи до того уровня, на котором полученные подзадачи просты. При этом онтологические знания есть активной основой процесса декомпозиции.

Онтологические знания, описывающие некоторое ПрП, можно разделить на следующие компонентные знания: типы входных и выходных данных, инструменты, операторы (человек или компьютерная программа) и операции (действие оператора или  $\text{J}_{\text{r}}\text{f}_{\text{f}}\text{e}_y \text{ a}_{\text{z}}$ ) [9].

Для реализации подхода необходимо разработать унифицированный язык представления онтологических знаний и инструментальную среду как набор специализированных и универсальных базовых операций, управляющих процессом решения. Необходимо также разработать  $\text{N}_{\text{h}}\text{j}_{\text{f}}\text{b} \text{ - } \text{j}_{\text{h}}\text{k}_{\text{e}}\text{v} \text{ a}_{\text{z}}$ , осуществляющий выбор средств и методов формирования структуры задачи на основе базовых операций.

Схема модели онтологии задач описывается кортежем

$$O^3 = \langle OZ^{\text{ПрП}}, M, PZ \rangle, \quad (4)$$

где  $OZ^{\text{ПрП}}$  – обобщенная задача проблемного пространства, состоящая из  $\text{J}$  задач, которые в свою очередь со-

стоят из  $q = \overline{1, Q}$  фрагментов каждая. Каждый фрагмент представлен процедурой, реализованной на множестве  $r = \overline{1, R}$  операций каждая;  $M$  – множество методов решения задач;  $PZ$  – решатель задач.

На рис. 4 представлена схема онтологии задач.

Одно из преимуществ онтологического подхода, в частности, иерархического представления состоит в том, что сложная задача большой размерности разбивается на последовательно решаемые группы задач малой размерности. Онтология задач в качестве понятий содержит типы решаемых задач, а отношения этой онтологии, как правило, специфицируют декомпозицию задач на подзадачи.

На рис. 5 представлен алгоритм методики проектирования онтологии ПдО и онтологии задач. Предполагается, что множество функций интерпретации в моделях онтологий тождественно множеству аксиом. На рисунке приняты следующие сокращения:

- КТ – контрольные точки;
- ПрП – проблемное пространство;
- ТС – толковый словарь;
- Т-О, Т-П – термины-объекты и термины-процессы;
- $X^O, R^O, F^O$  – конечные множества понятий, отношений и функций интерпретации онтологии объектов;
- $X^{\text{П}}, R^{\text{П}}, F^{\text{П}}$  – конечные множества понятий, отношений и функций интерпретации онтологии процессов;

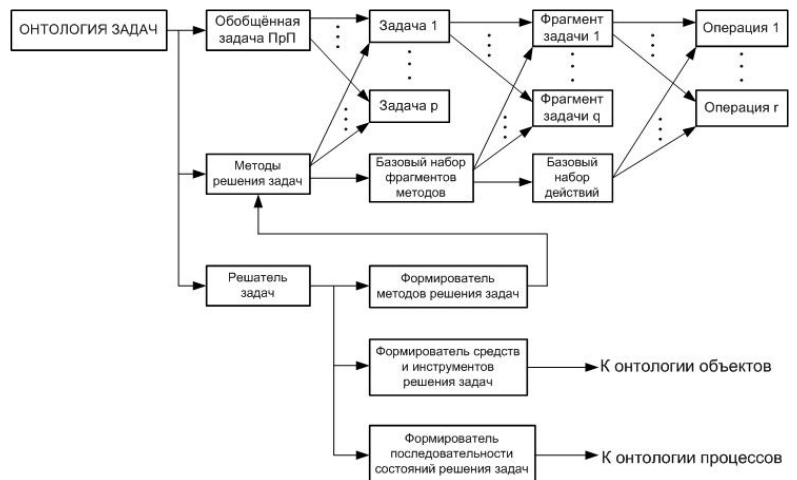


Рис. 4. Схема онтологии задач

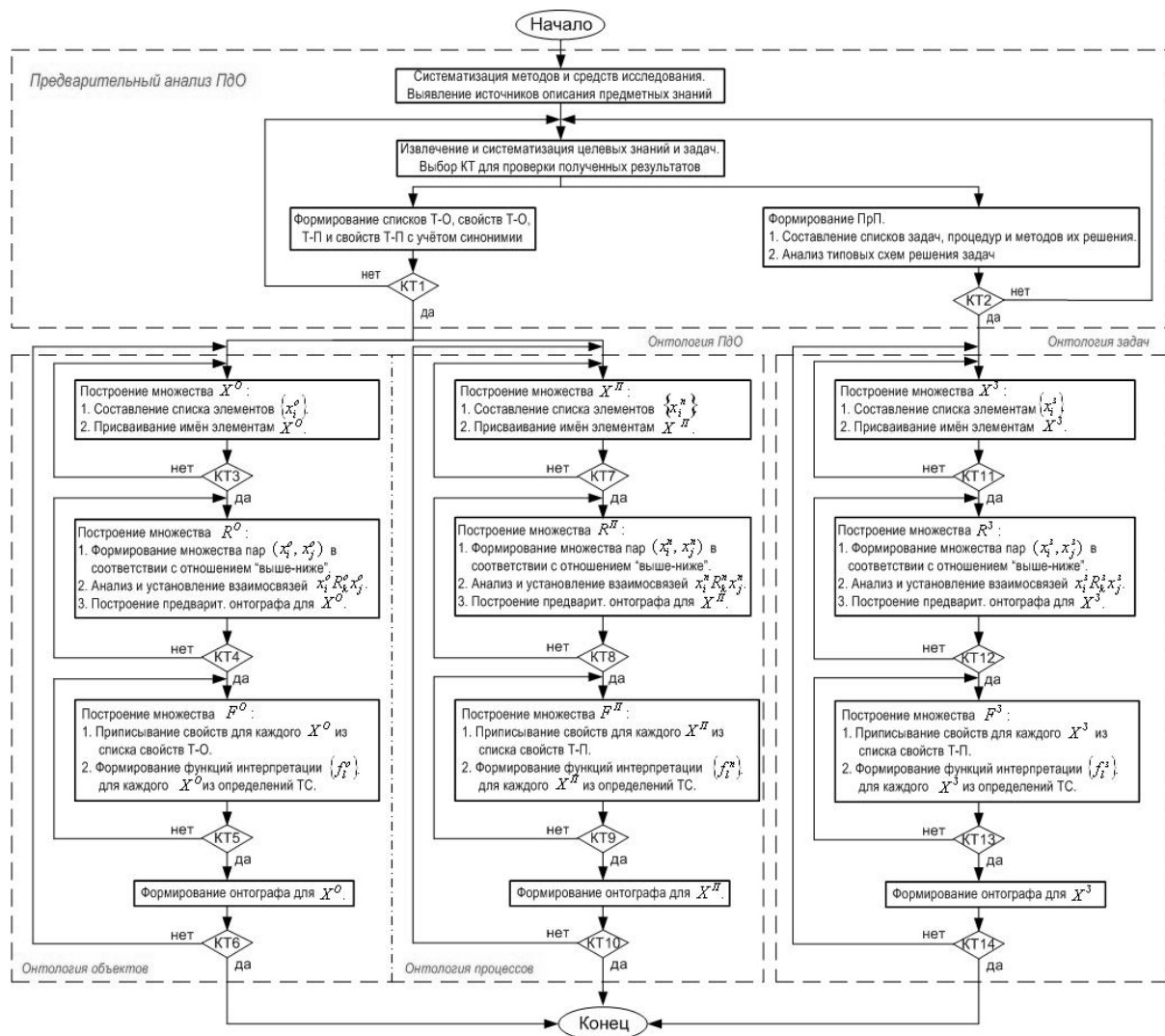


Рис. 5. Граф-схема алгоритма проектирования онтологий

–  $X^3, R^3, F^3$  – конечные множества понятий, отношений и функций интерпретации онтологии задач;

– КТ1 – анализ полноты сформированных списков терминов;

– КТ2 – анализ полноты сформированных списков задач, процедур и методов;

– КТ3 (КТ7, КТ11) – выполняется анализ «Список  $X^O (X^II, X^3)$  – исчерпан?»;

– КТ4 (КТ8, КТ12) – выполняется анализ «Отношения установлены для всех пар  $(x_i^O, x_j^O) [(x_i^i, x_j^i), (x_i^a, x_j^a)]$ ?»;

– КТ5 (КТ9, КТ13) – выполняется анализ «Функции интерпретации определены для всех  $X^O (X^II, X^3)$ ?»;

– КТ6 (КТ10, КТ14) – анализ полноты и непротиворечивости онтологии объектов (процессов, задач) экспертом.

Adexq\_gb. На основе анализа особенностей знаниеориентированных информационных систем предложен общий подход к построению онтологической базы знаний, функционирующей в составе онтологоуправляемой информационной системы, названный системно-онтологическим. Компонентами онтологической базы знаний являются онтология объектов, онтология процессов и онтология задач, для которых разработаны схемы моделей и методика проектирования. СОП и соответствующая технология его реализации предполагают

анализ ПДО, извлечение, представление и обработку предметных знаний. Вопрос автоматизации приобретения знаний, в частности из множества естественно-языковых объектов, по-прежнему актуален.

1. *Lhedhuc* словарь по искусственному интеллекту / Авт.-сост. А.Н. Аверкин, М.Г. Гаазе-Рапопорт, Д.А. Поспелов. – М.: Радио и связь, 1992. – 256 с.
2. *Object-oriented modeling and design* / J. Rumbaugh, M. Blaha, W. Premerlan et al. // Englewood Cliffs, 1991. – New Jersey: Prentice Hall. – 180 p.
3. *Ontology Development 101: A Guide to Creating Your First Ontology* / N.F. Noy, D.L. McGuinness // Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001. – [http://protege.stanford.edu/publications/ontology\\_development/ontolo-gy101.html](http://protege.stanford.edu/publications/ontology_development/ontolo-gy101.html).
4. *Gruber T.R.* A translation approach to portable ontology specifications // *Knowledge Acquisition*, 1993. – 5 (2). – P. 199–220.
5. *Nbehkhkndh*-методологические основания системных исследований: Системный анализ и системное моделирование. – М.: Наука, 1983. – 324 с.
6. *Ajmjhkdb* F.A., *Igdjhh* G.>. Системный анализ: проблемы, методология, приложения. – К.: Наук. думка, 2005. – 744 с.
7. *E\_lhu* и средства автоматизированного проектирования прикладной онтологии / Б.В. Добров, Н.В. Лукашевич, О.А. Невзорова и др. // *Изв. РАН. Теория и системы управления*. М.: 2004. – № 2. – С. 58–68.
8. *G\_ahjh* H.: . Онтолингвистические системы: методологические основы построения // Научная сессия МИФИ-2007. Сб. науч. тр. Т 3. Интеллектуальные системы и технологии. – М.: Наука, 2007. – С. 84–85.
9. *E\_lhbq\_kdb* комплекс по дисциплине «Современные проблемы науки» // *Методология анализа предметных знаний*. – [http://old.ulstu.ru/people/SOS-NIN/umk/Modern\\_Scientific\\_Problems/beloborodov/item\\_znan.htm](http://old.ulstu.ru/people/SOS-NIN/umk/Modern_Scientific_Problems/beloborodov/item_znan.htm)
10. *Iej* H., *Dmj* H. Міждисциплінарні наукові дослідження: оптимізація системно-інформаційної підтримки // *Вісн. НАН України*. – 2009. – № 3. – С. 14–25.
11. *Fjd* >.: ., *Fdhnwg* D. Методология структурного анализа и проектирования. – М.: МетаТехнология, 1993. – 239 с.
12. *E\_lhhehby* функционального моделирования IDEF0. РД IDEF0 – 2000 // Госстандарт России. – М. – 2000. – 75 с. – [www.nsu.ru/smk/files/idef.pdf](http://www.nsu.ru/smk/files/idef.pdf).
13. *IDEF5 Method Report*. – Armstrong Laboratory AL/HRGA Wright-Patterson Air Force Base, Ohio 45433. – 187 p. – [www.idef.com/pdf/Idef5.pdf](http://www.idef.com/pdf/Idef5.pdf)
14. *Jech* L.: ., *Ohjhr\_kdb* <N. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. – 384 с.
15. *Iejb* : .<, *Ydhe* X .K. Системная интеграция средств компьютерной техники. – Винница: УНІ-ВЕРСУМ-Вінниця, 2005. – 680 с.
16. *Gomez-Perez A., Fernandez-Lopez M., Corcho H.* Ontological engineering. – Springer, 2004. – 225 p.
17. *Desz* : .K., *Ijfv* B.E. Отношения между онтологиями предметных областей. Ч.1. // *Информационный анализ*, Вып. 1. – 2002. – 2. – С. 4–9.
18. *Iej* : .F ., *Ijadbg* >.<, *Kbfdh* D.<. Особенности проектирования модели и онтологии предметной области для поиска противоречий в правовых электронных библиотеках. – [www.inteltec.ru/publish/articles/textan/RCDL2004.shtml](http://www.inteltec.ru/publish/articles/textan/RCDL2004.shtml)
19. *Igh* N.B ., *Ymgbg* E.?. , *Iagbq\_gdh* <B. Логические модели интеллектуальных информационных систем. – К.: Наук. думка, 1999. – 397 с.
20. *hrcbeh* ?>.D. Понятие. – М.: МГУ, 1967. – 286 с.
21. *Be* X <. Логика: Учебник для вузов. – М.: Логос, 1997. – 272 с.
22. *Sowa J.F.* Conceptual Graphs as a universal knowledge representation // *Semantic Networks in Artificial Intelligence*, Spec. Issue of An Intern. J. Computers & Mathematics with Applications. (Ed. F. Lehmann), 1992. – 23, N 2–5, Part 1. – P. 75–95.
23. *SUO*, (2001), The IEEE Standard Upper Ontology web site, <http://suo.ieee.org>
24. *Niles I., Pease A.* Towards a Standard Upper Ontology. // *Proc. of the 2<sup>nd</sup> Intern. Conf. on Formal Ontology and Inform. Syst. (FOIS-2001)*, eds. C. Welty C, B. Smith, Ogunquit, Maine. – October, 17–19, 2001. – 8 p. – <http://home.earthlink.net/~adampease/professional/FOIS.pdf>
25. *The Mikrokosmos Ontology*. – <http://crl.nmsu.edu/Research/Projects/mikro/htmls/asis.paperhtml/node4.html>
26. *Iej* H.<, *Ijgdh* F.=. Модель категоріального рівня мовно-онтологічної картини світу // *Математичні машини і системи*. – 2006. – № 3. – С. 91–104.

Поступила 30.04.2009

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

© А.В. Палагин, Н.Г. Петренко, 2009