К ВОПРОСУ ВИЗУАЛИЗАЦИИ ОНТОГРАФОВ ПРИ РАЗРАБОТКЕ ОНТОЛОГИЙ ПРЕДМЕТНЫХ ДИСЦИПЛИН БОЛЬШОГО ОБЪЕМА

Палагин А.В., Петренко Н.Г., Величко В.Ю., Тихонов Ю.Л.

Аннотация*. В работе описана методика, предназначенная для повышения эффективности графического представления и выразительности онтологий предметных дисциплин (ПдД). Предложена система индексации, ориентированная на машинное представление и обработку онтографа ПдД, и оригинальный инструментарий, реализующий информационную технологию построения онтологии ПдД и ее онтографа, электронных курсов обучения по различным ПдД. Интерфейс содержит всплывающие подсказки, в которые включено имя понятия и время, необходимое на изложение описания данного понятия, что поможет преподавателю выбрать конкретное содержание учебного курса в соответствии с ограничениями "Рабочей программы" ПдД. Приведен пример построения, визуализации и индексации онтографа ПдД "Базы данных" (БД).*

Ключевые слова*: онтологический инжиниринг, представление и обработку онтографа*

ACM Classification Keywords: *I.2 ARTIFICIAL INTELLIGENCE - I.2.4 Knowledge Representation Formalisms and Methods, K.3 COMPUTERS AND EDUCATION - K.3.1 Computer Uses in Education*

Введение

В настоящее время e-learning развивается параллельно с World Wide Web, совершенствуя свои методики и формы обучения. Во многих странах, включая Украину, вводятся новые формы обучения, базирующиеся на компьютерных и телекоммуникационных технологиях, в том числе такие формы e-learning, как электронные курсы (ЭК).

E-learning и ЭК требуют методологических и нормативно-технологических разработок. Для ВУЗов, их кафедр и преподавателей e-learning и ЭК представляют новую, слабо обеспеченную методологически, нормативно и технологически, область педагогического знания и новые педагогические технологии. ЭК часто носят субъективный характер. Не обеспечено единое междисциплинарное пространство знаний. Отсутствуют библиотеки онтологичных баз знаний. Имеющиеся инструментальные средства разработки ЭК не учитывают междисциплинарных связей предметных дисциплин и не нацелены на уменьшение субъективности ЭК. Вышесказанное обуславливает необходимость разработки концепции создания онтолого-управляемых ЭК на основе онтологий ПдД, в которой должна быть определена методология поддержки разработки онтолого-управляемых ЭК на основе онтологий ПдД.

Такая методология в числе прочего должна включать [4]:

– перевод e-learning с уровня, основанного на информации на уровень, основанный на знаниях, в том числе онтологических;

– обеспечение высокой степени формализации представления онтологии ПдД и механизмы онтолого-управления как для процесса обучения, так и проверки усвоения знаний;

– обеспечение эффективной машинной обработки общеязыковых и предметных знаний ПдД на основе компьютерных онтологий (КО) ПдД;

– использование инструментальных средств поддержки автоматизированного построения онтологий ПдД;

– обеспечение строгой структуризации терминов и понятий ПдД.

К онтологическим аспектам относится круг вопросов, начиная от областей применения и до формального описания компонент компьютерных онтологий ПдД. Главный вектор исследований направлен на формализацию этапов построения, структурирования и представления знаний ПдД и интегрированного с ней проблемного пространства (эффективное сочетание лекционного и лабораторного материалов).

Онтология ПдД обладает когнитивностью, структурирует знания некоторой предметной области в соответствии с онтологическими принципами. Преподавателю онтология (в особенности компьютерная) облегчит разработку учебного курса, студенту улучшит восприятие материала и взаимосвязи предметных знаний. В силу этого можно ожидать значительного эффекта от применения онтологий в образовании в виде улучшения качества восприятия знаний в ПдД.

Автоматизированная разработка ЭК на основе лингво-семантического анализа больших объемов текстовой информации на первом этапе использует КО ПдД, которая может включать значительное число понятий. Формализованный онтограф ПдД и соответствующие инструментальные средства снижают затраты на разработку структуры и содержания ЭК. Однако онтограф может занимать значительный объем, затрудняющий цельное (когнитивное) восприятие ПдД, структуру знаний ПдД.

В данной работе рассмотрена методика по повышению эффективности графического представления и выразительности онтографов предметных дисциплин (ПдД).

Постановка задачи

Значимым практическим приложением онтологического инжиниринга является онтологический поход (ОнП) к проектированию и разработке ЭК. При таком подходе могут быть решены (и качественно улучшены решения) актуальные задачи в области e-learning [4]:

– автоматизированная разработка ЭК на основе лингво-семантического анализа больших объемов текстовой информации с помощью оригинальных инструментальных средств;

– структурирование и обработка терминов и понятий, содержащихся в ОУ в соответствии с утвержденной рабочей программой по заданной ПдД;

– существенное уменьшение трудоемкости составления (обновления) ЭК, в т.ч. новых ЭК.

В онтолого-классификационной схеме средств и методов искусственного интеллекта [1] ОнП трактуется как разновидность системного подхода, основанного на знаниях. ОнП обеспечивает эффективное проектирование компонент любой ЗОИС.

При разработке знание-ориентированной системы подготовки ЭК с учетом *онтологического подхода учитываются требования* [5]:

1. КО ПдД обеспечивают эффективную машинную обработку общеязыковых и предметных знаний ПдД.

2. В отличие от разработки ЭК традиционными методами (субъективного в своей основе), системно-онтологический подход [2] предполагает строгую структуризацию терминов и понятий. Проектирование *онтологии верхнего уровня* (ОВУ), представляющей собой онтологию домена предметных дисциплин кафедры, должно быть включено в общий алгоритм разработки ЭК.

3. Необходимо использование средств поддержки автоматизированного построения онтологий ПдД (методика, технология и программная реализация).

КО ПдД можно рассматривать как общезначимую, открытую базу знаний ПдД, представленную на общепринятом (формальном) языке спецификации знаний. Это позволяет выделить из множества разнообразных практических приложений КО [3] приложение с особым статусом – обучение. Создание "онтологизированных" электронных курсов (в том числе и дистанционных) представляется важной самостоятельной научно-технической проблемой [4]. При этом естественным образом находят своё решение такие проблемы как: открытости и закрытости систем обучения; единообразия представления понятийных структур обучения; автоматизированного построения новых курсов обучения с учётом накопленных преподавателями кафедры знаний и др.

Ожидается, что онтологии будут влиять на следующее поколение систем e-learning.

В учебном процессе гораздо эффективнее показать студенту онтограф ПдД, чем другими средствами описывать концепты ПдД и их связи. Онтология ПдД включает большое число понятий (как объектов, так и процессов). При визуализации онтологии онтограф может занимать значительный объем, затрудняющий восприятие отношений между понятиями ПдД, структуру знания ПдД. Это может помешать получить полный эффект от применения онтологического подхода в e-learning.

В связи с этим необходима методика, предназначенная для повышения эффективности графического представления и (когнитивной) выразительности онтографов ПдД.

Индексация онтографа

При построении КО с помощью оригинальных инструментальных средств возможно визуально представлять онтограф в компактном виде и организовать быстрый поиск по дереву понятий на основе эффективной индексации вершин онтографа. В частности, такие методы индексации широко используются в СУБД.

В работе предлагается система индексации, позволяющая сжать визуальное представление онтографа ПдД. В индексе каждого понятия в наглядной и удобочитаемой форме отражается уровень, связи между вершинами как по вертикали (связи между уровнями онтологии ПдД), так и по горизонтали (связи между понятиями-объектами и понятиями-процессами одного уровня представления).

Индекс имеет следующую структуру:

УN.n-УN1.n1-…-УNi.ni…-УNk.nk--п, где:

N – номер уровня, в котором располагается вершина онтографа, визуально представляющего онтологию;

N – номер понятия онтологии на уровне ***N***;

Ni.ni – номер уровня и номер понятия, с которым данное понятие связано;

n – признак процесса.

Например, индекс ***У5.2-У0.1-У3.5--n*** описывает второе понятие из уровня 5, связанное с первым понятием уровня 0 и с пятым понятием уровня 3.

Средства инструментальной поддержки

Одной из характерных черт знание-ориентированной системы разработки ЭК является обеспечение эффективной машинной обработки компьютерных онтологий ПдД.

Предложенная индексация ориентирована на машинное представление и обработку онтографа ПдД. Достигается единообразие представления понятийных структур ПдД. За счет этого при компьютерной обработке формального OWL–описания индексированного онтографа требуется меньший объем вычислений. В инструментарий, реализующий новую информационную технологию построения онтологии ПдД и ее онтографа, электронных курсов обучения по различным предметным дисциплинам, добавляется оригинальная программа визуализации индексированного онтографа ПдД. В качестве входных данных используется формальное описание индексированного онтографа ПдД на языке OWL. Интерфейс содержит всплывающие подсказки, в которые включено имя понятия, соответствующего вершине индексированного онтографа ПдД, и, при необходимости, объем описания данного понятия, что поможет преподавателю выбрать конкретное содержание учебного курса в соответствии с ограничениями “Рабочей программы” ПдД. Описание понятия содержится в файле “Определения онтологии”.

Блок-схема программы визуализации индексированного онтографа ПдД приведена на рис.1.

Рассмотрим пример индексации онтологии – индексации онтологии ПдД БД.

На рис. 2 представлен фрагмент онтографа онтологии ПдД БД.

На рис. 3 представлен фрагмент индексированного онтографа онтологии ПдД “Базы данных”, соответствующий аналогичному неиндексированному (рис. 2).

На рис. 5 представлен фрагмент индексированного онтографа онтологии ПдД БД, спроектированный в инструментальной среде Protégé, а на рис. 6 – фрагмент его OWL-описания.

Описание понятий индексированного онтографа онтологии ПдД БД содержится в файле “ОО”. На рис. 7 представлен фрагмент "ОО" для индексированного онтографа онтологии ПдД БД.

На рис. 8 представлен фрагмент формы программы визуализации индексированного онтографа ПдД. На сканах формы видны всплывающие подсказки, в которые включено имя понятия, соответствующего вершине индексированного онтографа ПдД. Программа разрабатывается на объектно-ориентированном языке JAVA, в соответствии с блок-схемой (рис. 1).

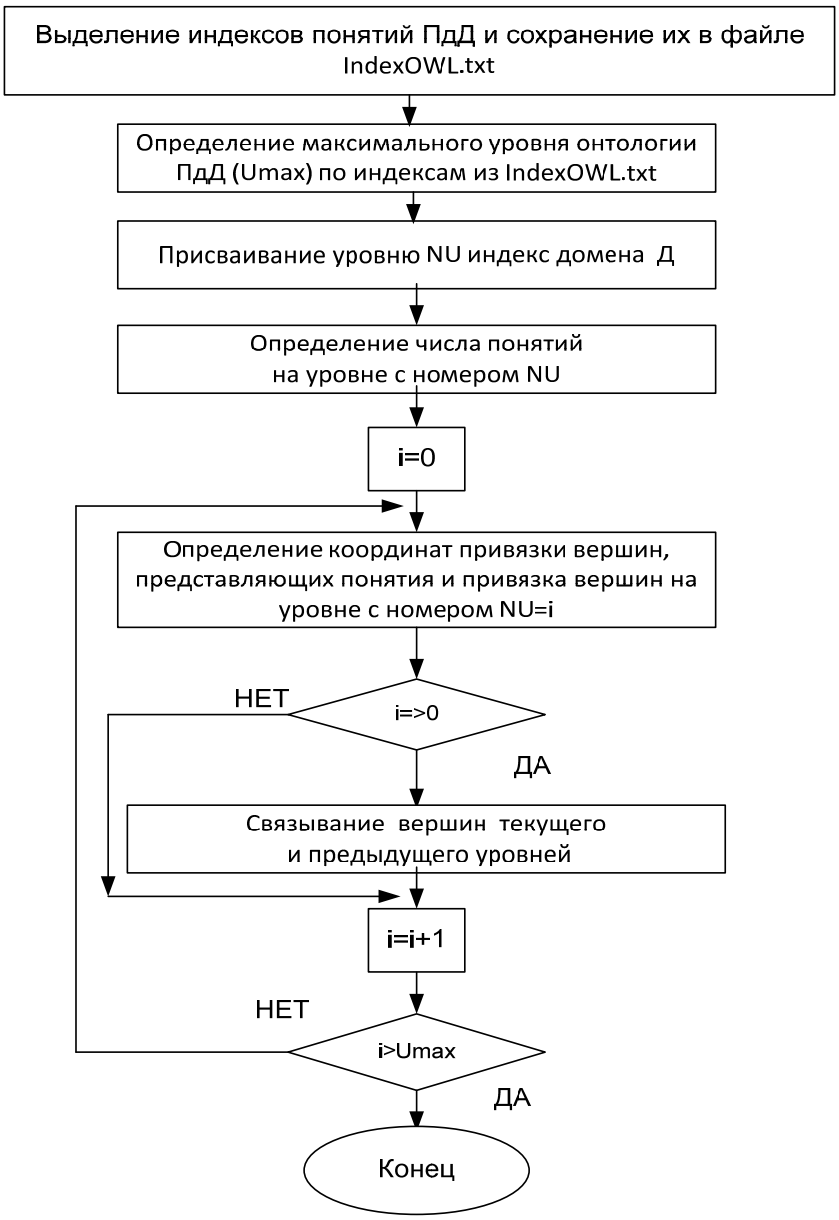

Рис. 1. Укрупненная блок-схема программы визуализации индексированного онтографа ПдД

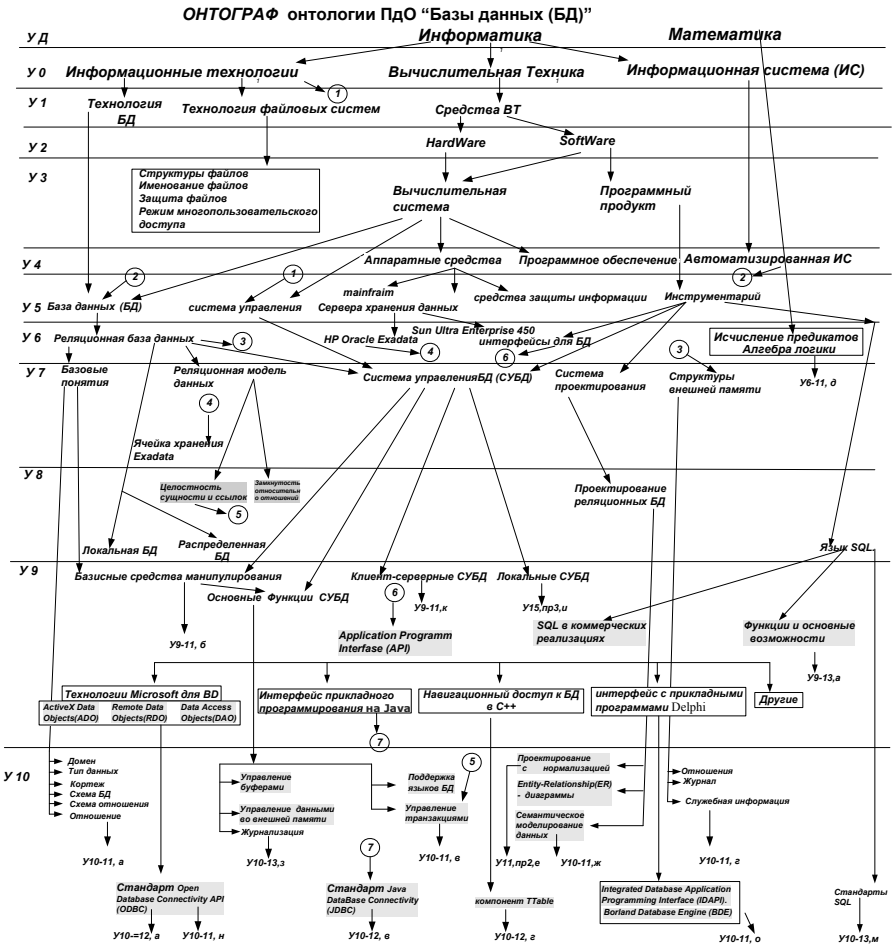

Рис.2 Фрагмент онтографа онтологии ПдД БД

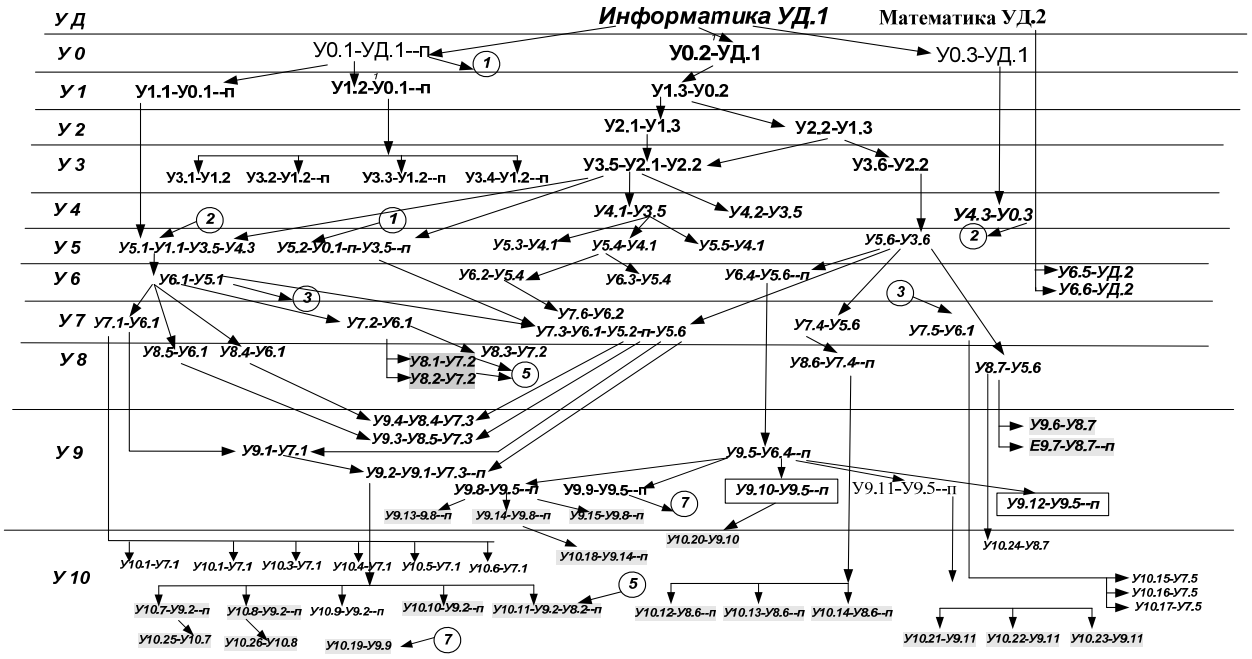

Рис. 3 Фрагмент индексированного онтографа онтологии ПдД БД

Компактная форма представления в индексированном онтографе онтологии ПдД “Базы данных” облегчает разделение объектов и процессов в визуальном представлении онтографа (рис. 4).

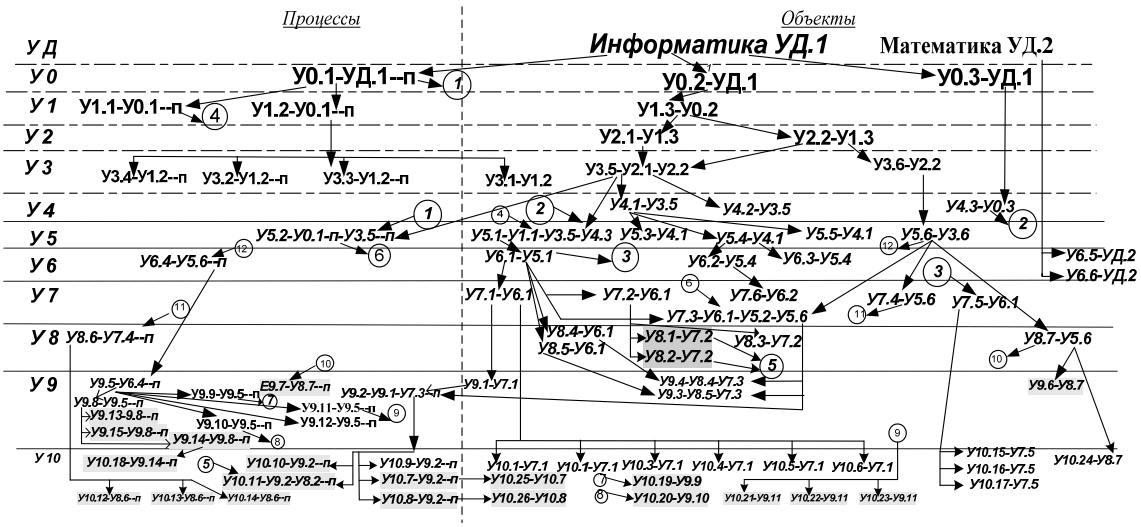

Рис.4. Фрагмент индексированного онтографа онтологии ПдД “Базы данных” с разделением понятий-объектов и понятий-процессов

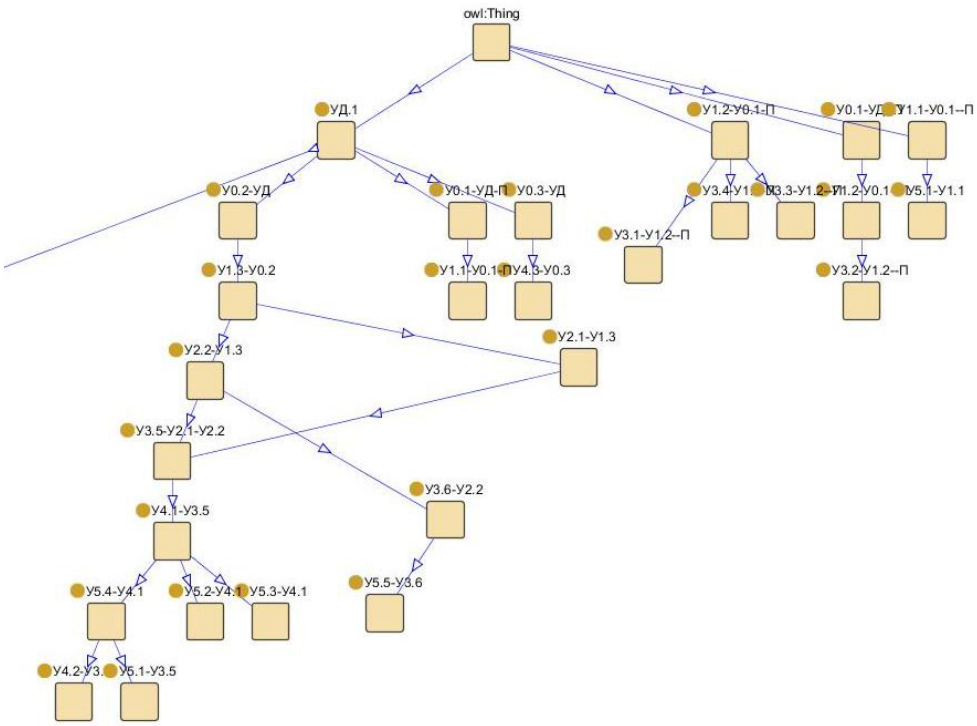

Рис.5 Фрагмент индексированного онтографа онтологии ПдД БД.

```
index_owl.owl - Lister
Файл Правка Вид Справка
  <?xml version="1.0" ?>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
    xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns="http://www.co-ode.org/ontologies/pizza/2005/10/18/pizza.owl#"
    xmlns:swrl="http://www.w3.org/2003/11/swrl#" xmlns:swrlb="http://www.w3.org/2003/11/swrlb#" xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xml:base="http://www.co-ode.org/ontologies/pizza/2005/10/18/pizza.owl">
  - <owl:Ontology rdf:about="">
      <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#string">version 1.3</owl:versionInfo>
      <rdfs:comment xml:lang="en">An example ontology that contains all constructs required for the various versions of the Pizza Tutorial run by Manchester
        University (see http://www.co-ode.org/resources/tutorials/)</rdfs:comment>
      <owl:imports rdf:resource="http://protege.stanford.edu/plugins/owl/protege" />
      <protege:defaultLanguage rdf:datatype="http://www.w3.org/2001/XMLSchema#string">en</protege:defaultLanguage>
    </owl:Ontology>
  - <owl:Class rdf:ID="У3.2-У1.2--П">
    - <rdfs:subClassOf>
        <owl:Class rdf:ID="У1.2-У0.1--П" />
      </rdfs:subClassOf>
    </owl:Class>
  - <owl:Class rdf:ID="У5.2-У4.1">
    - <rdfs:subClassOf>
        <owl:Class rdf:ID="У4.1-У3.5" />
      </rdfs:subClassOf>
    </owl:Class>
  - <owl:Class rdf:ID="У5.3-У4.1">
```


Рис. 6. Фрагмент OWL-описания индексированного онтографа онтологии ПдД БД.

№ вершины	Имя вершины (концепта)	Определение концепта
УД.1	Информатика	(Computer science, CS) – 1) научная дисциплина, изучающая структуру и общие свойства научной информации, а также закономерности всех процессов научной коммуникации; 2) большая область теоретических и прикладных знаний, связанных с получением, сохранением, преобразованием, пересылкой и использованием информации; 3) научное направление, что занимается изучением законов, методов и способов накопления, переработки и передачи информации с помощью компьютера и других технических средств; группа дисциплин, которые занимаются разными аспектами использования и разработки ЭВМ: прикладная математика, программирование, программное обеспечение, искусственный интеллект, архитектура ЭВМ, вычислительные сети.
У0.1-УД.1—п	Информационные технологии	(ИТ) – 1) приобретение, обработка, хранение и передача голосовой, графической, текстовой и числовой информации посредством основанной на микроэлектронной комбинации компьютера и телекоммуникационных связей; 2) практическая деятельность и прикладная наука, имеющие дело с данными и информацией. - information technology, IT — широкий класс ПдД, относящихся к технологиям управления и обработки данных, а также создания данных, в том числе, с применением вычислительной техники.
У0.3-УД.1	Информационная система	Это система, обеспечивающая ввод, хранение, поиск, обработку и вывод данных по запросам или регулярно необходимых данных.
У0.2-УД.1	Вычислительная техника	1) дисциплина, изучающая вычислительные машины, принципы их построения и использования. Включает исследование таких аспектов, как: программирование, информационные структуры,

№ вершины	Имя вершины (концепта)	Определение концепта
		разработка программного обеспечения, языки программирования, компиляторы и операционные системы, разработка аппаратных средств и тестирование, архитектура вычислительных систем, сети ЭВМ и сопряжение вычислительных машин, системный анализ и проектирование, теория информации, систем и вычислений, прикладная математика и электроника, вычислительные методы, методы машинной графики и приложения; 2) область техники, объединяющая средства автоматизации математических вычислений и обработки информации в различных областях человеческой деятельности.
У1.1-У0.1—п	Технология БД	Организация, принципы построения и функционирования современных систем управления базами данных, методы моделирования данных, принципы построения приложений баз данных.
У1.2-У0.1—п	Технология файловых систем	Термин файловая система (file system) используется для обозначения программной системы, управляющей файлами, и архива файлов, хранящегося во внешней памяти.
У1.3-У0.2	Средства ВТ	
У2.1-У1.3	HardWare	Аппара́тное обеспече́ние[1] (англ. hardware ˈha:dwɛə), жарг. «железо») — электронные и/или механические части вычислительного устройства (компьютер, ЭВМ, микроЭВМ и тд.), исключая его программное обеспечение и данные (информация, которую он хранит и обрабатывает).
У2.2-У1.3	SoftWare	ПО – совокупность программ системы обработки информации и программных документов, необходимых для эксплуатации этих программ (ГОСТ 19781-90)

Рис. 7. Фрагмент ОО для индексированного онтографа онтологии ПдД БД

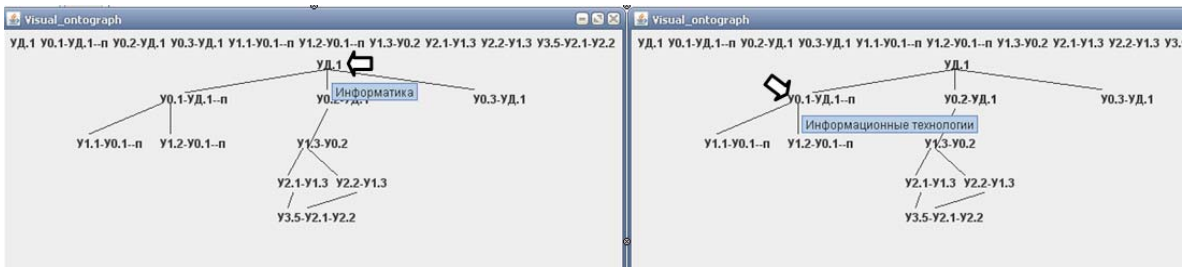

Рис. 8. Фрагмент формы программы визуализации индексированного онтографа ПдД

Для визуального просмотра, с использованием всплывающих подсказок, можно использовать вид онтографа, представленный на рис. 9. В этом случае индекс включается в подсказку.

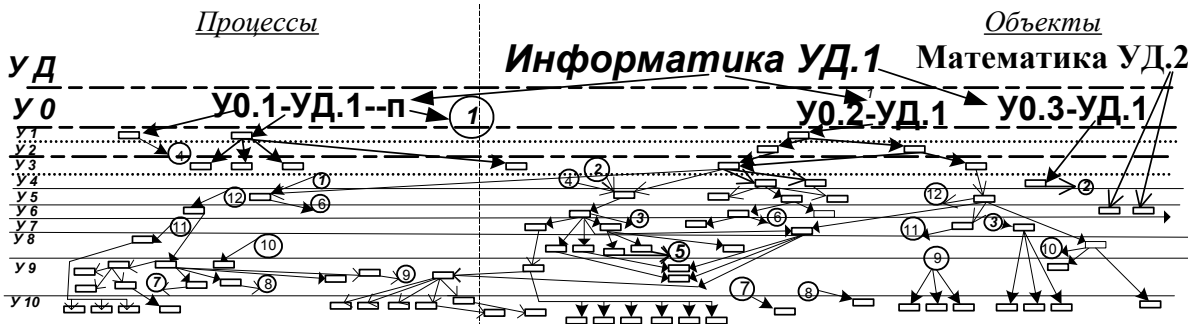

Рис. 9. Онтограф на основе всплывающих подсказок

Выводы

Предложенная система индексации, ориентированная на машинное представление и обработку онтографа ПдД является частью методики, предназначенной для повышения эффективности графического представления и выразительности онтологий предметных дисциплин (ПдД). Как результат подобной методики можно ожидать, что графические 3D-образы онтологий большого объема будут способны нести в себе в сжатой и одновременно с этим доступной для пользователя форме информацию достаточную для принятия адекватного решения. В частности, при разработке электронных курсов обучения по различным ПдД онтограф поможет преподавателю выбрать конкретное содержание учебного курса в соответствии с ограничениями “Рабочей программы”. Анимация графических 3D-образов онтологий обеспечит приемлемую навигацию по онтографу большого объема.

Развитие работ, связанных с индексацией вершин онтографов будут решены следующие задачи:

1. Автоматизация индексации онтографа с использованием OWL-описания онтологии ПдД.

2. Ускоренный поиск понятий по онтографу большого объема на основе разработанной системы индексации.

3. Разработка знание-ориентированного АРМ для ВУЗов Украины, с включением спецпроцессора обработки индексной формализованной онтологии и экрана большого размера.

ЛИТЕРАТУРА

1. Палагин А.В. Знание-ориентированные информационные системы с обработкой естественно-языковых объектов: основы методологии и архитектурно-структурная организация / А.В. Палагин, С.Л. Крывый, Н.Г. Петренко // УСиМ. – 2009. – № 3. – С.42–55.
2. Палагин А.В. Системно-онтологический анализ предметной области / А.В.Палагин, Н.Г.Петренко // УСиМ.–2009.– № 4. – С.3–14.
3. Башмаков А.И. Интеллектуальные информационные технологии: Учеб. Пособие / А.И.Башмаков, И.А Башмаков – М.: Изд. – во МГТУ им. Н.Э. Баумана, 2005. – 304 с.

4. Палагин А.В. К вопросу автоматизированного построения онтологии предметной дисциплины для электронных курсов обучения / А.В.Палагин, Н.Г.Петренко, Ю.Л.Тихонов, В.Ю Величко // Вісник східноукраїнського національного університету імені Володимира Даля – 2010. – № 4(150). – С.171–179.
5. Александр Палагин. Знание-ориентированные системы разработки электронных курсов обучения / Александр Палагин, Юрий Тихонов, Николай Петренко, Виталий Величко // Information Models of Knowledge ITHEA® / Krassimir Markov, Vitalii Velychko, Oleksy Voloshin (ed.) – Kiev, Ukraine – Sofia, Bulgaria, 2010 ISBN 978-954-16-0048-1. – PP. 304–313.

Информация об авторах

Палагин Александр Васильевич *– академик НАН Украины; заместитель директора Института кибернетики имени В.М. Глушкова НАН Украины; Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: palagin_a@ukr.net*

Основные области научных исследований: системная интеграция трансдисциплинарных научных знаний, онтологический инжиниринг

Тихонов Юрий Леонтьевич *– к.т.н., доц., доцент кафедры ИТС, Луганского национального университета им. Тараса Шевченко; e-mail: t2003i@mail.ru*

Основные области научных исследований: разработка онтологизированных электронных курсов

Петренко Николай Григорьевич *– к.т.н., старший научный сотрудник, Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: petrng@ukr.net*

Основные области научных исследований: методология и инструментальные средства автоматизированного проектирования онтологий предметных областей, системная интеграция трансдисциплинарных научных знаний

Величко Виталий Юрьевич *– к.т.н., старший научный сотрудник, Ин-т кибернетики*
им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: velychko@aduis.com.ua

Основные области научных исследований: индуктивный логический вывод, обработка естественно-языковых текстов.