

УДК 321.3

**В.В. Бова****МОДЕЛИРОВАНИЕ ОБЛАСТИ ЗНАНИЙ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ НЕПРЕРЫВНОГО ПРОФЕССИОНАЛЬНОГО ОБУЧЕНИЯ\***

*Работа посвящена рассмотрению метода моделирования области знаний в системе непрерывной обучения с помощью иерархической сетевой модели образовательного Web-контента и понятийно-тезисной модели представления знаний. Сведения об области знаний используются в процессах поддержки принятия решений при генерации персональной образовательной среды и автоматизированного тестирования знаний пользователей системы непрерывного профессионального обучения.*

*Методы принятия решений; адаптивные обучающие системы; модель предметной области; семантическая модель.*

**V.V. Bova****SIMULATION OF THE FIELD OF KNOWLEDGE IN DECISION SUPPORT SYSTEM FOR CONTINUOUS PROFESSIONAL**

*This paper is devoted to consideration of the method of modeling knowledge in a system of continuous learning through a hierarchical network model of educational Web-content and concepts and тезисной model of knowledge representation. Information about the knowledge used in decision support in the generation of personalized educational environment, and automated testing of knowledge of users of the system of continuing professional education.*

*Methods of decision-making; adaptive training systems; model of a subject domain; semantic model.*

**Введение.** В условиях стремительного развития средств распространения информации и усовершенствования технологий продуцирования Web-ресурсов большое значение имеет интенсивная интеллектуальная поддержка информационных ресурсов, расширяющих целевые для обучения области знаний при формировании персонального информационного пространства [1]. Невзирая на значительные достижения в области интеллектуальных систем обучения, средств контроля знаний и систем распространения информации в Интернет, актуальной остается задача разработки интегрированного системного подхода, который органично соединит методы сохранения и управления знаниями, индивидуализируемое представление образовательного Web-контента и методы автоматизации контроля знаний.

Проблема моделирования области знаний характеризуется рядом задач, которые имеют отношение к управлению знаниями на базе Интернет и возникают в контексте управления контентом, адаптации и персонализации контента, подготовки и сопровождения непрерывной профессионального обучения, генерации индивидуальных учебных курсов и автоматизированного контроля знаний. Адаптивность системы обучения определяется в применении динамических учебных курсов, в противовес статическим и в моделировании образовательного контента для обеспечения работы образовательных запросов при построении

\* Работа выполнена при поддержке: РФФИ (грант № 09-07-00318), г/б 12.8.08, г/б № 2.1.2.1652.

индивидуальной образовательной среды. Формирование образовательных запросов основывается на адаптивном выборе необходимой области знаний и определении в ней образовательного контента, который будет выдан пользователю для обучения в качестве индивидуального образовательного курса.

В данной статье описывается метод моделирования области знаний в системе непрерывной обучения с помощью иерархической сетевой модели образовательного Web-контента и понятийно-тезисной модели представления знаний. Сведения об области знаний используются в процессах поддержки принятия решений при генерации персональной образовательной среды и автоматизированного тестирования знаний пользователей системы непрерывного профессионального обучения.

**Сетевая модель образовательного Web-контента.** В основе модели Web-контента лежат две ключевые сущности: элемент контента и тематическая группа. Все элементы контента являются объектами, характеристики которых хранятся в таблице базы данных, а методы реализации – в классе, который описывается в ядре системы. Элементы контента называются Представлениями (Views), отвечающими за представление отдельного, логически завершенного, самостоятельного объекта контента. Каждый объект имеет уникальный символьный идентификатор, который задает пользователь системы при создании нового объекта. Таким образом, для идентификации отдельного элемента контента не нужны данные о его положении в структуре сайта. Поэтому при изменении места элемента в меню его Интернет-адрес (URL) не изменяется. Кроме того, во всех случаях использования объекта внутри системы для его идентификации используется лишь данный идентификатор, который дает преимущества и способствует реализации семантического управления контентом.

**Иерархическая структура содержимого.** Все объекты контента упорядочиваются в виде иерархической структуры. Объекты первого уровня иерархии представляют основные разделы Web-ресурса. Каждый из объектов может иметь произвольное количество дочерних объектов, которые, в свою очередь, также могут иметь собственные дочерние объекты и т.д. Говоря об иерархической структуре контента, следует отдельно задать характерную для Web-контента конструкцию типа «список - элементы списка». В зависимости от типа объекта, его основная часть генерируется автоматически. Так объект контента типа «Список» (list) являет собой перечень вступительных блоков (аннотаций) элементов с ссылками на дочерние объекты, то есть на элементы списка. С помощью такой функциональности удобно формировать такие разделы Web-ресурса как база статей, архив новостей и тому подобное. Таким образом, конструкцию «список - элементы списка» применяется для реализации раздела Web-ресурса, который содержит большое количество однотипных элементов контента.

**Сетевая структура содержимого – контент как семантическая сеть.** Между объектами контента поддерживаются ассоциативные отношения. Идея ассоциативного отношения выражается в том, что определенная страница логически связана по смыслу с другой страницей или элементом контента. Ассоциативные страницы могут указываться явно в специальном поле объекта контента. Ассоциативные элементы также добавляются автоматически в случае, если в тексте страницы встречается ссылка на другую страницу Web-ресурса в виде аннотированной ссылки или в виде вступительного блока с приглашением читать дальше. Еще один специфический тип отношения, который устанавливается между объектами, – отношения *идентичности*, когда один объект играет роль *псевдонима* другого элемента, то есть является его копией или клоном. Кроме того, проектируется функ-

ция *тематической каталогизации* объектов контента. Ее идея заключается в формировании *тематической структуры* альтернативной к центральной структуре объектов с целью реализации адаптивного тематико-ориентированного представления Web-контента.

Для построения многопредметного Web-портала для организации непрерывного профессионального обучения, используются блоки контента.

**Семантический блок** – это множественное число элементов контента, имеющих логическое и структурное единство и единственный источник происхождения, например, одно авторство или тема. Сюда можно отнести готовые статические курсы, контент которых загружен в систему. Физически блок контента является некоторой ветвью в дереве контента. При этом, обозначив вершину этого дерева как блок, можно свести всех его потомков к данному блоку. Пример семантических блоков изображен на рис. 1. Элементы, обозначенные как блок становятся частью нового семантического блока вместе со всеми потомками.

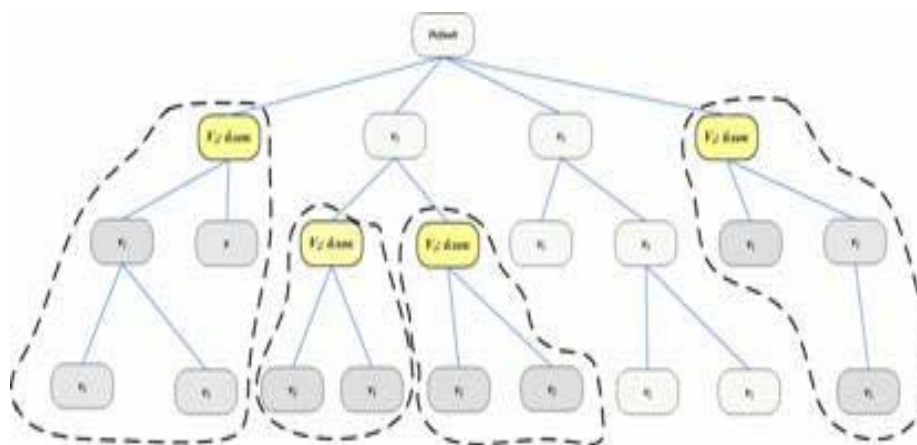


Рис. 1. Пример семантических блоков

Таким образом модель Web-контента состоит из двух древовидных структур: структура элементов контента и структура тематических групп. Каждый элемент контента может содержать произвольное количество тематических групп. Дерево контента указывает на физическое расположение объектов, это основная навигационная модель контента Web-ресурса. Тогда как дерево тематических групп носит семантический характер и определяет иерархию предметных областей. Схематически модель контента изображена на рис. 2.

Множественное число элементов контента:  $V=\{v_1, \dots, v_n\}$ . Иерархическая структура контента: каждый элемент может иметь дочерние элементы:  $Ch:v>2^V$ . Каждый элемент имеет родительский элемент:  $F:v>V$ . Сетевая структура контента: каждый элемент имеет совокупность связанных с ним элементов:  $N:V>2^V$ . Множественное число  $G$  указывает на тематические или ассоциативные группы, в которых могут принимать участие элементы контента:  $G=\{g_1, \dots, g_{ng}\}$ . Организация тематических групп происходит в иерархической структуре. Каждый элемент контента может принимать участие в произвольном количестве тематических групп:  $GV:V>2^G$ . На основе тематических групп определяются связи ассоциативности между элементами.

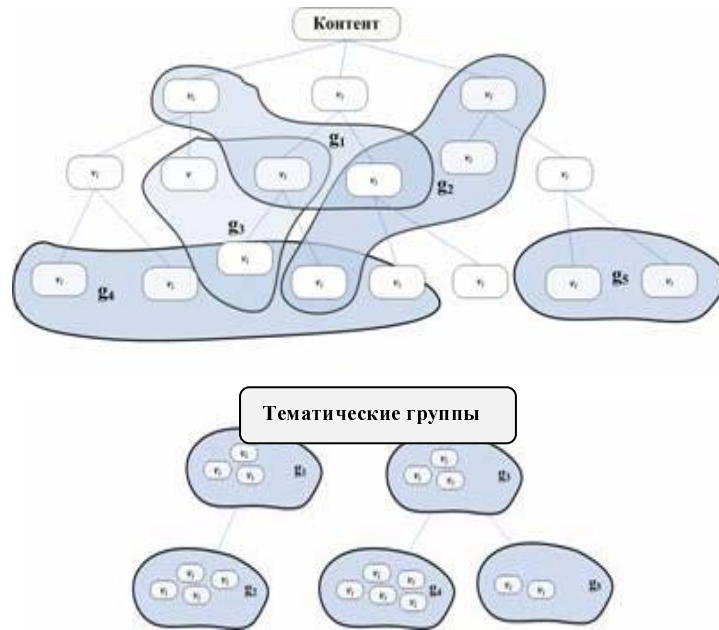


Рис. 2. Схематическое изображение модели: дерево контента и дерево тематических групп

**Понятийно-тезисная модель (ПТМ).** В образовательном процессе можно выделить три составляющие: передача учебной информации, социальная составляющая, которая включает общение между участниками образовательного процесса, практические занятия [2]. Самая первая миссия дистанционного образования – упрощение и оптимизация передачи учебной информации от носителя знаний к обучающемуся. Исторически сложилось, что в традиционных формах обучения начальной формой передачи информации является вербальный способ (преподаватель читает лекцию). В дистанционном обучении – основной формой передачи учебной информации становится текст, а точнее гипертекст и мультимедиа содержимое. То есть основным транспортным средством или же «протоколом» передачи знаний является текст, в котором содержится учебная информация. Можно сказать, что предложенная понятийно-тезисная модель представления знаний для образовательных целей имеет текстуальный характер. Именно в этом заключается принципиальное отличие ПТМ от других подходов к формализации знаний, традиционных для инженерии знаний.

Модель формализации дидактических текстов основывается на подходе разбора текста. Важный взнос в дидактический подход вносит *герменевтика* [3]. Так как эта наука занимается толкованием текста, ее положения, можно применить для толкования текста в понимании его формализации в компьютерной системе. Такие понятия, как контекст, ключевые слова в тексте, основная мысль автора должны помочь формализовать знания, которые находятся в тексте для их последующего представления и применения для проверки знаний обучающихся. В отличие от учебника, компьютерная система должна обеспечить адекватное к знаниям студента представление учебного материала. Здесь неотвратимая линейность классического печатного текста должна быть преодолена. В этом вопросе должно помочь понятие гипертекста.

Путем аналитического пересмотра текста и специального визуального интерфейса из текста экспертом выделяются и заносятся в базу знаний семантические сущности. Этот процесс можно сравнить с индексацией. Информация, которая хранится в ПТ-базе дает возможность делать выводы о том, какое понятие является предметом обсуждения в том или другом тексте. Также ПТМ дает возможность генерировать тесты для проверки усвоения студентом образовательного контента.

Основными семантическими сущностями ПТМ являются *понятие* и *тезисы*. Понятие выражает предмет знаний, который обсуждается в том или другом фрагменте учебного материала. Тезис – это некоторая ведомость или утверждение о понятии [4]. Каждому понятию отвечает собственный набор тезисов, извлеченных из текста, которые характеризуют данное понятие. Множественное число понятий  $C = \{c_1, c_{n2}\}$ . Множественное число тезисов:  $T = \{t_1, t_{n1}\}$ . Множественное число элементов контента  $V = \{v_1, v_{n3}\}$ . Каждый элемент контента  $v_n$  может иметь произвольное количество тезисов  $t_j$ :  $TV: V > 2^T$ . Каждый тезис  $t_j$  принадлежит одному элементу контента  $v_n$ .  $VT: T > V$ . Каждое понятие  $c_k$  может иметь произвольное количество тезисов  $t_n$ . Принадлежность тезисов понятиям задается отображением:  $CT: T > C$ . Каждое понятие имеет некоторый набор тезисов, который задается отображением:  $TC: C > 2^T$ .

**Интеграция моделей.** Комплексная модель образовательного контента представляет собой интеграцию рассмотренных моделей. При этом взаимная интеграция моделей порождает дополнительную семантику, которая описывает область знаний каждой из моделей и сведения об ассоциативности элементов контента. Под *областью знаний* в компьютерной системе понимается некоторое множественное число ассоциативных элементов контента, определенного в конкретной области знаний. Учитывая характер заданий, которые возникают в системе управления знаниями для непрерывного обучения, актуальным ставится вопрос определения области знаний для данного элемента контента. Получая на входе некоторый Web-ресурс системы, необходимо определить множественное число ассоциативного контента, который будем называть областью знаний определенного уровня детализации. Увеличивая объем такой области знаний, расширяется взгляд на предметную область, касаясь также других тем, и получается новое множество элементов контента. Принцип определения и последующего расширения области знаний элемента контента изображен на рис. 3. Выделяем разные области ассоциативности для элементов контента: *ближайший круг* тематически ассоциативных элементов контента; *умеренный круг* (углубляющий) тематически ассоциативных элементов контента; *широкий круг* (обобщающий) тематически ассоциативных элементов контента.

**Методика моделирования области знаний.** Проблема моделирования области знаний сводится к нахождению ассоциативных элементов контента данного Web-ресурса. Источники ассоциативности элемента контента по порядку значимости:

1. Элементы связаны общими понятиями ПТМ.
2. Бинарные связки между элементами контента.
3. Элементы-члены той же группы, к которой принадлежит данный элемент.
4. Элементы того же семантического блока контента.
5. Элементы из дочерних групп.
6. Элементы из родительских групп.
7. Иерархические связки в дереве контента: дочерние элементы, родительский элемент.

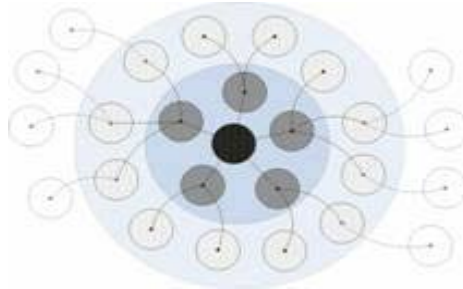


Рис. 3. Принцип определения и последующего расширения области знаний элемента

Кроме определения совокупности ассоциативных элементов, необходимо также упорядочить их внутри каждого из источников по степени, или рангу ассоциативности, к элементу, относительно которого происходит моделирование.

1. Элементы связаны общими понятиями ПТМ: ранг элементов прямо пропорционален количеству общих с исходным элементом понятий.

2. Бинарные связи между элементами контента: ранг всех элементов одинаков.

3. Элементы-члены той же группы, к которой принадлежит данный элемент: ранг элемента прямо пропорционален релевантности элементов его группы.

4. Элементы того же семантического блока контента. Здесь ранг распределяется следующим образом: наибольший – у первого дочернего элемента. Далее последовательно идут следующие дочерние элементы (с минимальным отрывом в ранге). Следующим по рангу является родительский элемент (при условии, что он является частью семантического блока).

5. Элементы из дочерних групп: ранг элемента прямо пропорционален релевантности элементов его группы.

6. Элементы из родительских групп: ранг элемента прямо пропорционален релевантности элементов его группы.

На рис. 4. представлено последовательное определение ассоциативных элементов с учетом иерархии групп (см. рис. 2).

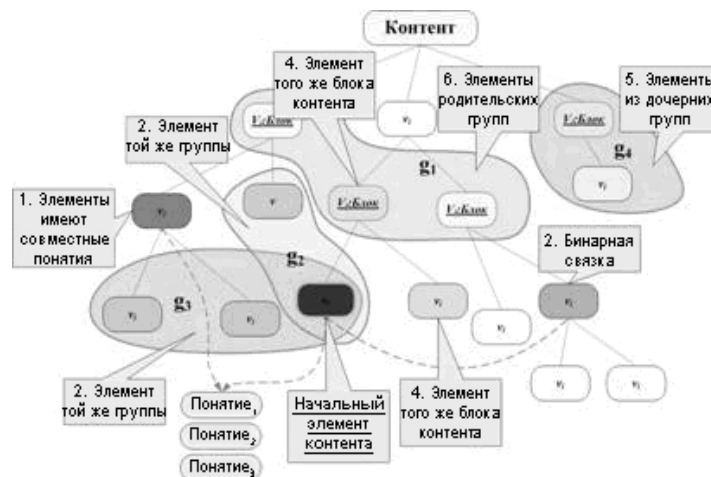


Рис. 4. Декомпозиция последовательного определения ассоциативных элементов области знаний

**Заключение.** Моделирование области знаний предоставляет возможность поддержать такие процессы организации непрерывного профессионального обучения через Интернет, как генерация персональной образовательной среды и автоматизированное тестирование.

Понятийно-тезисная модель предусматривает однозначную связь семантических данных с учебным материалом. Это дает возможность гибкого управления учебным процессом и процессом контроля знаний и их адаптации. В данном подходе предусмотрено построение дополнительных ссылок, организация навигации по виртуальной области знаний и генерации индивидуального учебного курса на портале дистанционного обучения. В случае тестирования на базе ПТМ, технология определения дополнительных ассоциативных элементов дает возможность использовать дополнительные семантические элементы ПТМ из других источников и генерировать тест и варианты ответов даже при недостаточном количестве понятий и тезисов в основном контенте. В действительности система не рассуждает знаниями из предметной области, однако относительно отдельных высказываний или утверждений, которые описывают понятие и воспринимаются системой «как есть», она владеет информацией о том, где они содержатся в общей структуре учебного материала.

Сам процесс формирования семантических данных происходит путем осмысленного чтения учебного текста и выделения из него понятий и тезисов с помощью интерфейса пользователя.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Демьянов А.В. Системная организация принятия управленческих решений в обучении // Известия вузов. Поволжский регион. Технические науки. – Пенза: Инф.-изд. центр ПГУ, 2006, №6. – С. 178-190.
2. Поспелов Д.А. Логико-лингвистические модели в системах управления. – М.: Энергоатомиздат, 1999.
3. Мысин М.Н. Использование информационных технологий в процессе формирования профессиональных компетенций будущего специалиста. – Самара.: Изд-во Самарского ун-та, 2004. – 194 с.
4. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2000.

**Бова Виктория Викторовна**

Технологический институт федерального государственного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: Vbova@mail.ru.

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 8(8634)371-651.

Кафедра систем автоматизированного проектирования; старший преподаватель.

**Bova Viktoriy Viktorovna**

Taganrog Institute of Technology – Federal State-Owned Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: Vbova@mail.ru.

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: 8(8634)371-651.

The Department of Computer Aided Design; senior teacher.