

## КОНСТРУИРОВАНИЕ КУРСОВ НА ОСНОВЕ СЕМАНТИЧЕСКОЙ СЕТИ НАУЧНЫХ ПОНЯТИЙ УЧЕБНОЙ ДИСЦИПЛИНЫ ДЛЯ ИНТЕРАКТИВНЫХ ОБУЧАЮЩИХ СИСТЕМ

Лутошкина Н.В.

*ФГБОУ ВО «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева», Красноярск, e-mail: lutoshkinanv@list.ru*

Цифровизация образования – одна из основных задач программы развития цифровой экономики. В статье рассмотрен подход к проектированию интерактивного курса на основе анализа научных терминов учебной дисциплины. Цель исследования заключается в разработке подхода к построению модели образовательного контента электронного обучающего курса, обеспечивающей формализованное представление учебного материала дисциплины. В результате анализа строятся две онтологии, связанные гиперссылками: онтология научных понятий и онтология учебных материалов, а также матрица смежности ациклического графа понятий учебной дисциплины. При проведении исследования использовались следующие методы: анализ научно-методической литературы по вопросам разработки оценочных материалов, методы логического анализа и теории графов, а также сравнительно-сопоставительный анализ психолого-педагогической литературы. В исследовании применялись методы обобщения и систематизации материала, анализа и синтеза теоретических аспектов изучаемой темы. Особенностью авторского подхода выступает структурирование предметной области в виде последовательности учебных объектов образовательного контента, изучаемых в определенном порядке. Описанная модель построения образовательного контента учебной дисциплины отличается от известных тем, что на основе логических методов анализа понятий построен граф понятий учебной дисциплины, на основе которого систематизируется контент учебного курса.

Ключевые слова: глоссарий, онтология, семантическая сеть, дерево понятий, образовательный контент, дистанционные курсы.

## DESIGNING COURSES BASED ON THE SEMANTIC NETWORK OF SCIENTIFIC CONCEPTS OF THE ACADEMIC DISCIPLINE FOR INTERACTIVE LEARNING SYSTEMS

Lutoshkina N.V.

*Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, e-mail: lutoshkinanv@list.ru*

Digitalization of education is one of the main objectives of the program for the development of the digital economy. The article considers an approach to designing an interactive course based on the analysis of scientific terms of the discipline. The purpose of the study is to develop an approach to building a model of educational content of an e-learning course that provides a formalized representation of the educational material of the discipline. As a result of the analysis, two ontologies are constructed: linked by hyperlinks, the ontology of scientific concepts and the ontology of educational materials, as well as the adjacency matrix of the acyclic graph of the concepts of the academic discipline. The following methods were used in the study: analysis of scientific and methodological literature on the development of evaluation materials, methods of logical analysis and graph theory, as well as comparative analysis of psychological and pedagogical literature. The study used methods of generalization and systematization of the material, analysis and synthesis of theoretical aspects of the topic under study. A special feature of the author's approach is the structuring of the subject area in the form of a sequence of educational objects of educational content, studied in a certain order. The described model of constructing the educational content of an academic discipline differs from the well-known ones in that a graph of the concepts of an academic discipline is constructed on the basis of logical methods of analyzing concepts, on the basis of which the content of the educational course is systematized.

Keywords: glossary, ontology, semantic network, concept tree, educational content, distance courses.

Современные образовательные процессы неразрывно связаны с использованием информационных технологий. Развитие современных средств информационных технологий,

уровень обеспечиваемых ими возможностей возросли настолько, что появилась возможность для решения задачи комплексной информатизации сферы образования. Одним из приоритетных направлений в этой области является создание интегрированной информационно-образовательной среды, объединяющей подсистемы различного масштаба и принадлежности и представляющей собой интеллектуальную систему, организованную на базе сети Интернет.

Одним из ключевых факторов, позволяющих производить объединение с минимальными издержками, является приведение компонент интерактивных образовательных систем к единым стандартам. Такими стандартами, предназначенными для систем электронного обучения, являются SCORM, IMS AF, IMS CC, RUS LOM [1].

Современные обучающие системы представляют собой интеллектуальные системы, основанные на парадигме обработки знаний. Термины составляют содержание знаний, в любой образовательной технологии необходимо владение системой понятий. Состав учебной системы знаний дисциплины образуют научные понятия, которые были включены в программу учебной дисциплины.

Цель исследования заключается в разработке подхода к построению модели образовательного контента электронного обучающего курса, обеспечивающей формализованное представление учебного материала дисциплины в условиях цифровизации образования.

### **Материал и методы исследования**

При исследовании был проведен анализ научно-методической литературы, касающейся вопросов разработки оценочных материалов, были применены методы логического анализа и теории графов, а также сравнительно-сопоставительный анализ психолого-педагогической литературы. В исследовании использовались методы обобщения и систематизации материала, анализа и синтеза теоретических аспектов изучаемой темы.

### **Результаты исследования и их обсуждение**

Как показывает практика, система понятий представляет собой не менее половины объема материала большинства учебных курсов. При этом усвоение оставшейся части материала, которая в основном содержит изложение структуры и функции системы знаний, а также примеров, невозможно без освоения понятийной составляющей дисциплины. Поэтому актуальной задачей, стоящей перед обучающей системой, представляется обеспечение средств для облегчения изучения и для контроля усвоения понятий, включенных в предметную область дисциплины курса. В большинстве современных интерактивных обучающих систем используется простейшее из подобных средств – глоссарий. Глоссарий

предназначен для работы с системой понятий учебного курса и может применяться не только для обучения, но и для контроля усвоения знаний [2].

Обучение, как и все процессы познания, требует систематизации знаний и, по возможности, целостного охвата всего предмета изучения. Этим целям в рамках построения учебного курса может служить семантическая сеть понятий, отображенная в виде графа, поскольку он является наиболее естественным для человеческого мозга видом структуры данных.

Действительно, система понятий может быть отображена в виде ориентированного графа, при этом каждое понятие отображается узлом, а ассоциативные связи между ними – ребрами графа. Если при этом граф не будет циклическим, то такой граф-дерево может отображать иерархию системы понятий предметной области дисциплины, а значит, его можно рассматривать как онтологию. Такая онтология может быть использована как основа для структуры базы знаний учебного курса, особенно при дистанционном обучении. Любой интерактивный курс, по сути, является набором модулей, которые заключают в себе множество обучающих ресурсов и средств контроля процесса усвоения материала. Поэтому онтология учебной дисциплины может быть исходной структурой для построения систем обучения и тестирования в рамках интерактивных курсов.

Таким образом, построение семантической сети понятий дисциплины как одного из видов онтологии является весьма актуальным для построения учебного курса, а также для успешного усвоения знаний, поскольку облегчает систематизацию изучаемых понятий [3].

Как показала практика, процесс построения онтологий является весьма трудоемкой процедурой [4].

Для упрощения и ускорения процесса создания онтологий, а именно семантических сетей, была предложена процедура автоматизации выделения семантических связей между изучаемыми понятиями на основе глоссария дисциплины с помощью специально разработанной программы. Программа позволяет построить семантическую сеть в виде направленного графа, узлы которого представляют изучаемые понятия, а ребра обозначают наличие между ними связи «определяемое понятие – определяющее понятие» [5].

В данном случае глоссарием называется словарь узкоспециализированных понятий, составляющих предметную область изучаемой дисциплины, содержащий их толкование, комментарии и ключевые примеры. При этом стоит отметить, что не всякий глоссарий может быть эффективно обработан для получения семантической сети – для качественной трансляции глоссарий должен удовлетворять, как минимум, следующим критериям:

- 1) понятие не может объясняться только через самое себя;

- 2) два понятия не могут быть связаны несколькими существенно разнородными отношениями и, желательно, не должны быть связаны взаимно обратными отношениями;
- 3) имеется одно ключевое понятие, через которое, напрямую или косвенно, определяются остальные понятия.

Также стоит отметить, что структура, правила составления и форма представления информационно-поисковых тезаурусов и глоссариев на русском языке регламентированы стандартом ГОСТ 7.25-2001 [2].

Поскольку каждый интерактивный обучающий курс содержит в себе, как минимум, список изучаемых понятий и их определение, систематизация контента учебного курса, основанная на анализе глоссария, доступна практически для любого интерактивного курса. Также весьма значимым достоинством данного метода является сжатая форма представления знаний в глоссарии, что позволяет избежать анализа большого объема данных, особенно всего образовательного контента.

В качестве входных данных в предлагаемом подходе используется простое текстовое представление глоссария изучаемой дисциплины — текстовый файл, содержащий набор определений в формате «определяемое понятие – определение», расположенных каждое в своем отдельном абзаце. Количество и структура предложений в определении, а также регистр букв и спецсимволы не имеют значения. Строки не в формате «определяемое понятие – определение» игнорируются.

Множество узлов графа, т.е. концептов составляемой семантической сети, формируются из выделенных в глоссарии понятий, которым было дано определение. Наличие связей между понятиями, которые будут отображаться в виде ребер графа, выявляется путем поиска упоминаний обозначенных концептов в тексте определения других концептов. В соответствии с этим строятся направленные ребра графа, от определяющего концепта к определяемому. При этом считается, что любой концепт, встречающийся в определении понятия, является определяющим для данного понятия независимо от используемой словоформы (рода, числа, падежа, лица, времени). Для этого в тексте определения отыскиваются не полные исходные словоформы концептов, а только основа слов – стемы, которые выделяются с помощью алгоритма Портера, модифицированного для русского языка.

Алгоритм стемминга, или стеммер Портера, использует морфологические особенности словообразования естественных языков и убирает суффиксы и окончания путем последовательного применения небольшого набора правил. Большим достоинством выбранного алгоритма является отсутствие необходимости в базе данных слов и, следовательно, применимость в случае узкоспециализированных, редких слов и неологизмов.

Кроме этого, исходный стеммер Портера для русского языка был дополнительно модифицирован таким образом, чтобы учитывался ряд суффиксов, используемых в образовании новых производных понятий, в результате чего была снижена частота ложного обнаружения связей.

Общая схема предлагаемого подхода выглядит следующим образом.

1. Сначала выявляется множество всех понятий, определяемых в предоставленном глоссарии: строятся вершины графа будущей семантической сети.

2. Далее для каждого понятия выделяются стемы, которые будут выявляться в тексте определений. Для понятий, состоящих из нескольких слов, выделяется последовательность соответствующих стемов.

3. Текст каждого определения обрабатывается стеммером для получения строки, состоящей из стемов всех слов, которые определяют данное понятие, с сохранением их порядка.

4. В полученных строках проводится поиск подстрок, содержащих стемы понятий, входящих в рассматриваемое множество: по наличию совпадений строятся направленные ребра графа, отображающие связи «определяемое понятие – определяющее понятие» семантической сети.

5. Осуществляются верификация и правка графа семантической сети пользователем.

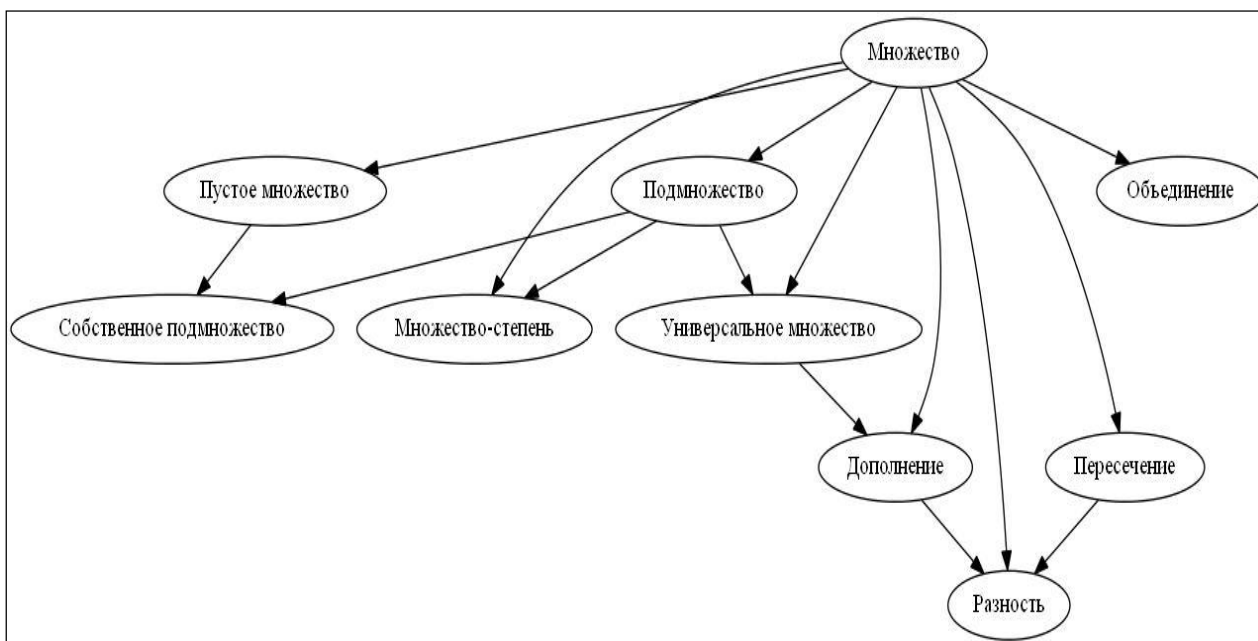
Конечным результатом обработки текста глоссария должен стать ориентированный граф-дерево, который отображает иерархическую семантическую сеть понятий, т.е. онтологию. Однако часто определения излишне ссылаются друг на друга, и тогда полученный граф может быть циклическим. При этом выделить из участвующих в цикле понятий наиболее базовое, через которое будут определены все остальные, практически невозможно в автоматическом режиме, поэтому требуется вмешательство пользователя.

Стоит также заметить, что у нескольких различных терминов могут получиться одинаковые стемы (стемы-омонимы). В данном случае связи не могут быть выделены, но в дальнейшем пользователь сможет установить их вручную, выстраивая связи между конфликтными определениями и остальными понятиями. Также может существовать термин, который способен представлять собой словосочетание, причем одно из слов (или некоторая подпоследовательность слов) имеет одинаковый стем с другим концептом (один концепт содержит в себе другой). При нахождении в определении подобных концептов включенные в него концепты не считаются определяющими (если только они не будут найдены отдельно в другом месте).

Для решения обозначенных выше проблем требуется участие пользователя, играющего при построении обучающего курса роль эксперта. Для этого ему предоставляются обработанный текст глоссария с выделенными проблемными определениями и промежуточный вариант графа понятий, визуализированного с помощью широко распространенного пакета Graphviz программой dot. Кроме этого, у пользователя имеется возможность править текст самого глоссария, добавляя или убирая ссылки на понятия внутри определений. Для удобства этого процесса текст глоссария дополнительно размечен, а фрагменты со стемами выявленных понятий помечены отдельно. Также программа отображает имеющиеся в графе циклы и даже может предложить ребра, удаление которых позволит избавиться от циклов. Для определения таких ребер-кандидатов программа рекурсивно пытается удалить по одному ребру, удаление которого максимально снижает полное число ребер, принадлежащих циклам.

После проверки и корректировки результата работы программы пользователь получает древовидный ориентированный граф концептов, т.е. семантическую сеть понятий изучаемой дисциплины. В дальнейшем она может быть использована для построения обучающего курса, включая анализ и структуризацию контента, а также систематизацию контроля усвоения материала.

На рисунке представлена иерархическая семантическая сеть понятий, полученная на основе анализа текста глоссария [1]. Понятия расположены сверху вниз от более общих к более частным.



*Графовое представление семантической сети понятий, сгенерированной на основе анализа текста глоссария курса*

Данная программа может использоваться при формировании курсов для:

- 1) построения семантической сети изучаемой дисциплины;
- 2) нахождения междисциплинарных связей (при использовании в качестве входного файла объединения нескольких глоссариев);
- 3) выделения подграфов при оценке сложности модулей рабочей программы учебной дисциплины;
- 4) вычисления характеристики важности концептов, на основе которой должны формироваться требования к банку тестовых заданий (для процентного содержания тестовых заданий с заданным концептом);
- 5) вычисления характеристики сложности концептов, на основе которой должна вычисляться сложность генерируемых тестов;
- 6) создания междисциплинарных тестов (на основе построения семантической сети объединенных глоссариев).

При конструировании интерактивного обучающего курса «Дискретная математика» для направления 15.03.04 были разработаны онтологии курсов. Онтология курса содержит онтологию описания структуры курса и предметную онтологию для описания понятий курса. Онтология описания структуры курса содержит гиперссылку на онтологию понятий курса. Каждый концепт онтологии понятий курса имеет гиперссылку на соответствующую статью глоссария курса. Онтология понятий курса разбита на части в соответствии с количеством зачетных единиц рабочей программы дисциплины. Студенты имеют возможность использовать ее как целиком, так и по частям. Все онтологии были встроены в курс СДО Moodle, их можно просматривать как рисунки графов, а также работать с ними, используя плагин Jambalay а в Protege [6].

Обучение требует постоянного контроля. На начальном этапе применяется формирующее оценивание. Формирующее оценивание призвано развивать интерес студента к прохождению курса и обязательно включает в себя развернутую обратную связь (комментарий) в случае выбора неверного ответа [7]. Такое тестирование также позволяет студенту узнать, что же он не понял из учебного материала, и повторить этот учебный материал перед прохождением итогового тестирования.

Система Moodle позволяет автоматически проверять и оценивать уровни понимания и умения использования понятийно-терминологического аппарата при помощи лекций или тестов. Элемент СДО Moodle «Лекция» представляет собой последовательность порций учебного материала и вопросов к ним; по сути, направлен не только на обучение, но и на контроль. Настройки «Лекции» позволяют предусматривать различные траектории ее прохождения в зависимости от уровня подготовленности и обучаемости студента. Тесты не

предусматривают включения учебного материала, но могут быть использованы как обучающие (с многократным количеством попыток), так и как проверочные в зависимости от целей использования. Проверочные тесты генерируются как случайная выборка из соответствующего тематике категории (раздела) банка тестовых заданий. Количество категорий банка тестовых заданий соответствует количеству контролируемых тем учебной дисциплины. Количество тестовых заданий в категории пропорционально количеству терминов, относящихся к теме категории. Уровень владения учебным материалом, умение его анализировать, систематизировать контролируются посредством использования элемента «Задание». Отчеты по выполненным заданиям студенты отправляют на сайт для проверки преподавателем. Работа проверяется и в случае неверного решения отправляется студенту для исправления ошибок. Итоговое тестирование направлено на оценку усвоения дисциплины. По результатам прохождения итогового тестирования определяется балл студента, с которым он допускается к экзамену.

**Выводы.** Особенностью авторского подхода выступает структурирование предметной области в виде последовательности учебных объектов образовательного контента, изучаемых в определенном порядке. Описанная модель построения образовательного контента учебной дисциплины отличается от известных тем, что на основе логических методов анализа понятий построен ациклический граф понятий учебной дисциплины, на основе которого систематизируется контент учебного курса. Кроме вышесказанного, матрицу смежности построенного графа можно использовать для междисциплинарных связей.

### Список литературы

1. Попов Д.И., Попова Е.Д., Певцов К.С. Обзор стандартов и спецификаций в электронном обучении и тестировании // Инновационные методы и средства оценки качества образования: материалы V Всероссийской научно-методической конференции. Московский государственный университет печати. [Электронный ресурс]. URL: <https://www.ast-centre.ru/books/favorites/279> (дата обращения: 24.04.2021).
2. ГОСТ 7.25-2001. Группа Т62. Межгосударственный стандарт. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления. Введен взамен ГОСТ 7.25-80 (СТ СЭВ 174-85) СИБИД. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления (с Изменением N 1). Дата введения 2002-07-01. М.: ИПК Издательство стандартов, 2001. 23 с.



3. Гущин Ю.В. Интерактивные методы обучения в высшей школе // Психологический журнал международного университета природы, общества и человека «Дубна». 2012. № 2. С. 1-18.
4. Лутошкина Н.В., Мурашова Л.М. Систематизация мультимедийного контента электронного курса на основе онтологии предметной области // Современные проблемы науки и образования. 2013. № 6. [Электронный ресурс]. URL: <http://science-education.ru/ru/article/view?id=11480> (дата обращения: 28.04.2021).
5. Кочеткова Т.О., Карнаухова О.А. Адаптивная образовательная стратегия обучения математике студентов в электронной среде // Вестник КГПУ им. В.П. Астафьева. 2018. № 2 (44). С. 50-56.
6. Галлямова О.Н. Технология профессиональной подготовки специалистов техносферной безопасности в вузе с использованием интерактивного подхода к обучению // Ученые записки университета им. П.Ф. Лесгафта. 2012. № 3 (85). С. 58-61.
7. Уваров А.Ю., Фрумин И.Д. Трудности и перспективы цифровой трансформации образования. М.: Издательский дом Высшей школы экономики, 2019. 342 с. DOI: 10.1007/s11423-019-09645-88.