

УДК 004:37

ОНТОЛОГИЯ ОРГАНИЗАЦИИ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА В ЗАДАЧАХ ПОИСКА И СОРТИРОВКИ

Спиваковский А.В., Осипова Н.В.

Херсонский государственный университет

Онтологии являются ключевой технологией семантической обработки знаний. Рассматривается методология использования онтологии для организации вычислительного эксперимента в задачах поиска и сортировки при изучении курса «Основы алгоритмизации и программирования».

Ключевые слова: онтология, интегрированная среда, вычислительный эксперимент.

1 Введение

Анализ современного состояния программных средств, приобретения знаний и поддержки учебной деятельности при изучении курса «Основы алгоритмизации и программирования» позволяет выявить две группы проблем.

- Методологические проблемы: слабая проработанность теоретических аспектов процессов извлечения знаний (философские, психологические, педагогические и другие аспекты); трудности в формировании логико-алгоритмического мышления, навыков формализованного описания поставленных задач, умений строить информационные модели, формировании знаний основных алгоритмических структур и умений применять эти знания для построения алгоритмов решения задач по их информационным моделям; проведении анализа эффективности разработанного алгоритма.
- Технологические проблемы: недостаток систем поддержки исследовательской деятельности, их узкой направленности (зависимость от платформы, языка реализации, ограничений), жесткость программных средств, их низкая адаптивность; слабые графические возможности, недостаточный учет когнитивных и эргономических факторов.

Интегрированная среда изучения курса «Основы алгоритмизации и программирования» (<http://weboar.ksu.ks.ua/>) позволяет вместе с изучением теоретического материала проводить вычислительный эксперимент для изучения эффективности алгоритмов поиска и сортировки. Такой подход к содержанию усиливает исследовательскую деятельность студентов, фундаментальную предметную подготовку будущих специалистов, за счет формально логического отображения причинно-следственных связей и, как следствие, влияет на мотивацию студентов [7]. Новое поколение систем управления знаниями основано на применении концептуальных онтологических моделей.

2 Постановка проблемы

Изменения в образовании на основе технологий электронного обучения (e learning) предполагают реализацию принципов самоорганизации, саморегулирования и самоконтроля открытых систем. Модель открытого образования исходит из открытости мира, процессов познания и образования человека. Основу виртуальной образовательной среды составляют создание и использование интегрированных сред обучения (ИСО), объединяющих знания научной дисциплины на основе принципов построения систем управления знаниями. ИСО обеспечивает интеграцию, накопление и поддержку, а также организацию доступа к знаниям виртуальной образовательной среды, что позволяет:

- объединить различные источники информации по дисциплине, специальностям и участникам образовательного процесса (преподавателям и студентам) в рамках единой системы;
- обеспечить постоянное развитие системы за счет обновления теоретического знания и непрерывного накопления нового опыта, полученного преподавателями и студентами в ходе учебного процесса;

- предоставлять релевантную решаемой задаче информацию каждому из участников образовательного процесса в соответствии с его знаниями, предпочтениями и потребностями.

Для повышения качества учебного процесса представляется необходимым интегрировать разработку учебных курсов путем координации выполняемых работ на основе методологии управления знаниями, базирующейся на объектном или компонентном подходе с онтологическим описанием концептуальной модели используемых знаний.

В объектном представлении знаний каждый учебный объект представляет законченный семантический фрагмент знаний, имеющий самостоятельное значение. Из отдельных объектов могут конфигурироваться конкретные учебные последовательности, соответствующие особенностям обучающихся с учетом региональных и отраслевых потребностей обучения. В результате происходит переход от больших негибких курсов к множествам многократно используемых объектов обучения (RLO – Reusable Learning Object), доступных для поиска и включения в конкретную учебную последовательность. Объектная методология формирования обучающих курсов нашла развитие в деятельности международного консорциума IMS Global Learning Consortium, разрабатывающего систему базовых стандартов в этой области.

Все объекты знаний располагаются в специальных хранилищах – репозиториях и собираются в учебный материал курса в момент исполнения учебного процесса конкретного обучающегося. Таким образом, реализуется компонентная методология создания обучающих последовательностей учебных курсов на основе интегрированного пространства знаний.

Для интеграции знаний в едином пространстве и последующего выделения компонентов учебных курсов (учебных объектов) требуется единое концептуальное описание знаний с помощью онтологий. Для организации интегрированного пространства знаний создаются: предметная онтология, отражающая виды деятельности независимо от того, кому и как они преподаются, и онтология обучения, формализующая структуру процесса обучения под углом зрения конкретных специальностей и форм обучения, а также репозиторий учебных объектов.

3 Анализ предыдущих публикаций

Онтологическое моделирование в области информатики (computer science) интенсивно развивается и применяется в таких областях как исследования в области искусственного интеллекта, представления знаний, обработки естественного языка, в Семантическом Вебе, в инженерии информационных систем и многих других. Разрабатываются и применяются разнообразные языки представления онтологий и их расширения для семантического описания конкретных артефактов, особенно в ИТ. Разработаны и развиваются онтологии в разнообразных предметных областях науки [1-3].

4 Основная часть

Для того чтобы определить структуру взаимосвязей между элементами знаний, необходимо выделить (абстрагировать) понятия из содержания элементов знаний (документов, опыта обучаемых) и структурировать (организовать) их формальным способом, путем задания взаимосвязей между этими понятиями. Одним из наиболее разработанных в настоящее время способов описания знаний в виде множества понятий и взаимоотношений между ними являются онтологии.

Онтологию можно определить как знаковую систему

$O = \{C, R, P, H, F, G, L, A\}$, где C – набор понятий онтологии, причем для каждого понятия $c \in C$ в онтологии существует, по крайней мере, одно утверждение; R – обозначает бинарный характер отношений между понятиями онтологии, фиксирующие пары области применения (domain)/области значений (range), то есть пары (D, R) с $D, R \in C$; P – множество свойств понятий; H – фиксирует таксономический характер отношений (связей), при котором понятия онтологии связаны неререфлексивными, ациклическими, транзитивными отношениями $H \subset C \times C$. Выражение $H(C_1, C_2)$ означает, что понятие C_1 является подпонятием (производным понятием) C_2 ; F и G – функции ссылок такие, что

F: FLC→2C и G: FLR→2R, то есть F и G связывают наборы лексических единиц {L_j} LC наборами понятий и отношений, на которые они соответственно ссылаются в данной онтологии, при этом одна лексическая единица может ссылаться на несколько понятий или отношений и одно понятие или отношение может ссылаться на несколько лексических единиц; A – набор аксиом онтологии [].

Использование онтологий для описания модели знаний позволяет использовать формальные языки для их описания, например, такие как, RDFS и OWL. А также воспользоваться строгим аппаратом дискретной логики для выполнения операций логического вывода над понятиями и метаданными онтологии. Использование онтологий позволяет интерпретировать понятия (и соответствующие им термины) как специалистами, так и компьютерными программами – системами управления знаниями (СУЗ).

Классификацию онтологий производят по следующим основным признакам: степени зависимости от конкретной задачи или предметной области; уровню детализации аксиоматизации; «природе» предметной области [4-5, 9].

В соответствии с этими признаками онтологии подразделяются на следующие виды:

- верхнего уровня – описывают наиболее общие концепты (пространство, время, материя, объект и т. д.), которые независимы от конкретной проблемы или области (например, проект СУС – база знаний понятий окружающего мира);
- онтологии предметной области – содержат понятия конкретной предметной области и множество интерпретаций этих знаний и отношений (например, общецелевая онтология UNSPSC, предоставляющая терминологию товаров и услуг);
- онтологии задач – используются конкретными программными продуктами и содержат термины, которые применяются при разработке программных систем, выполняющих конкретную задачу, например, онтология транспортных потоков товаров SCTG. При этом они используют специализацию терминов, представленных в онтологиях верхнего уровня, в частности UNSPSC;
- прикладные онтологии – описывают концепты, которые зависят как от конкретной предметной области, так и от задач, которые в них решаются. Концепты в таких онтологиях часто соответствуют ролям, которые играют объекты в предметной области в процессе выполнения определенной деятельности (например, онтология TOVE, которая обеспечивает построение интегрированной модели предметной области, состоящей из онтологий: операций, состояний, времени, организации, ресурсов, продуктов, сервиса, производства, цены и количества).

В качестве основных функций онтологий при использовании их в качестве составляющих информационной системы выделяют:

- эффективное компактное представление системы знаний конкретной предметной области на базе современных информационных технологий (спецификация и концептуализация);
- поиск информации в системе знаний (справочные и обучающие системы);
- постановка и решение прикладных задач в заданной предметной области (научные исследования, проектирование объектов новой техники и технологий и др.);
- развитие системы и получение новых знаний либо упорядочение существующих, проверка их непротиворечивости, коррекция категориального дерева;
- поиск информации в Интернет (онтология выступает посредником между пользователем и поисковой системой).

В жизненном цикле знаний можно выделить три основных этапа:

1. выявление потребности в знаниях;
2. производство (создание) знаний;
3. потребление (утилизация) знаний.

В различных моделях, эти этапы включают разное количество процессов. С учетом того, что знания являются конкурентным преимуществом, необходима целенаправленная деятельность по стимулированию накопления и использования знаний студентами.

Управление знаниями при использовании интегрированных сред обучения заключается в управлении жизненным циклом знаний за счет ведения целенаправленной деятельности по следующим направлениям (процессам в жизненном цикле знаний):

идентификация существующих знаний; создание новых знаний; хранение знаний; распространение знаний; и стимулирование использования знаний.

Технология конфигурирования процесса электронного обучения с использованием интегрированных сред обучения на основе онтологий предметной области и обучения, репозитория учебных объектов включает следующие технологические операции:

- формирование и обновление модели обучающегося;
- генерация сценария обучения в соответствии с моделью обучающегося;
- обучение в соответствии со сценарием.

Интегрированные среды обучения являются сложными системами, которые характеризуются иерархической упорядоченностью компонентов (подсистем) различных уровней.

На рисунке 1 приведена диаграмма компонентов интегрированной среды изучения курса «Основы алгоритмизации и программирования» [6], которая включает следующие модули:

- Система персонификации – обеспечивает регистрацию пользователей, предоставление им прав доступа (администратора, учителя и ученика), защиту данных системы от несанкционированного доступа.
- Учебное пособие – современное мультимедийное гипертекстовое приложение, построенное в виде структурированной коллекции тем и алгоритмов, содержание которого соответствует программе курса основы алгоритмизации и программирования для ВУЗов.
- Библиотека лекций – презентации лекционных материалов – электронные документы в формате .ppt , разработаны для поддержки лекционного курса в соответствии с разделами учебного пособия.
- Библиотека задач – построена в виде системы задач, классифицированных по темам и уровням сложности, в соответствии с разделами учебного пособия.

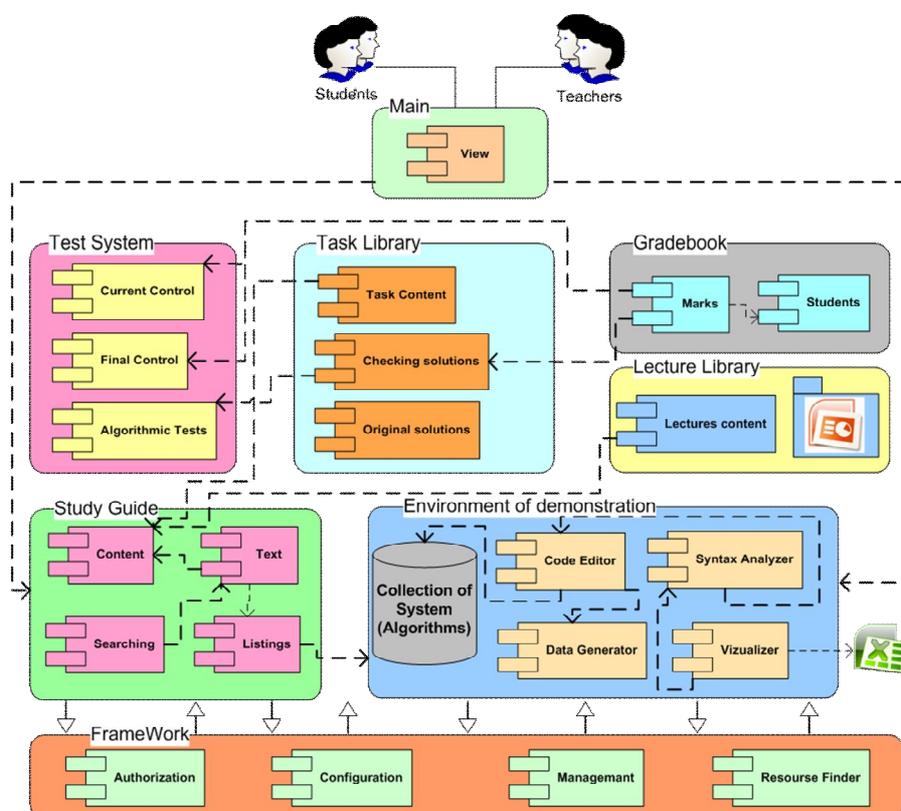


Рис. 1. Диаграмма компонентов интегрированной среды изучения курса «Основы алгоритмизации и программирования»

- Среда демонстрации программ используется при выполнении практических и лабораторных работ для наглядной демонстрации работы алгоритмов и организации

вычислительного эксперимента по изучению эффективности алгоритмов поиска и сортировки.

- Система текущего и итогового контроля знаний включает как традиционную систему тестов, так и модуль алгоритмического тестирования (тестирования программного кода), который является специфическим для данной предметной области.
- Система редактирования контента предоставляет возможность осуществлять разные траектории обучения, учитывая уровневую и профильную дифференциацию.
- Электронный журнал позволяет осуществлять мониторинг усвоения студентами знаний и умений по составлению алгоритмов.

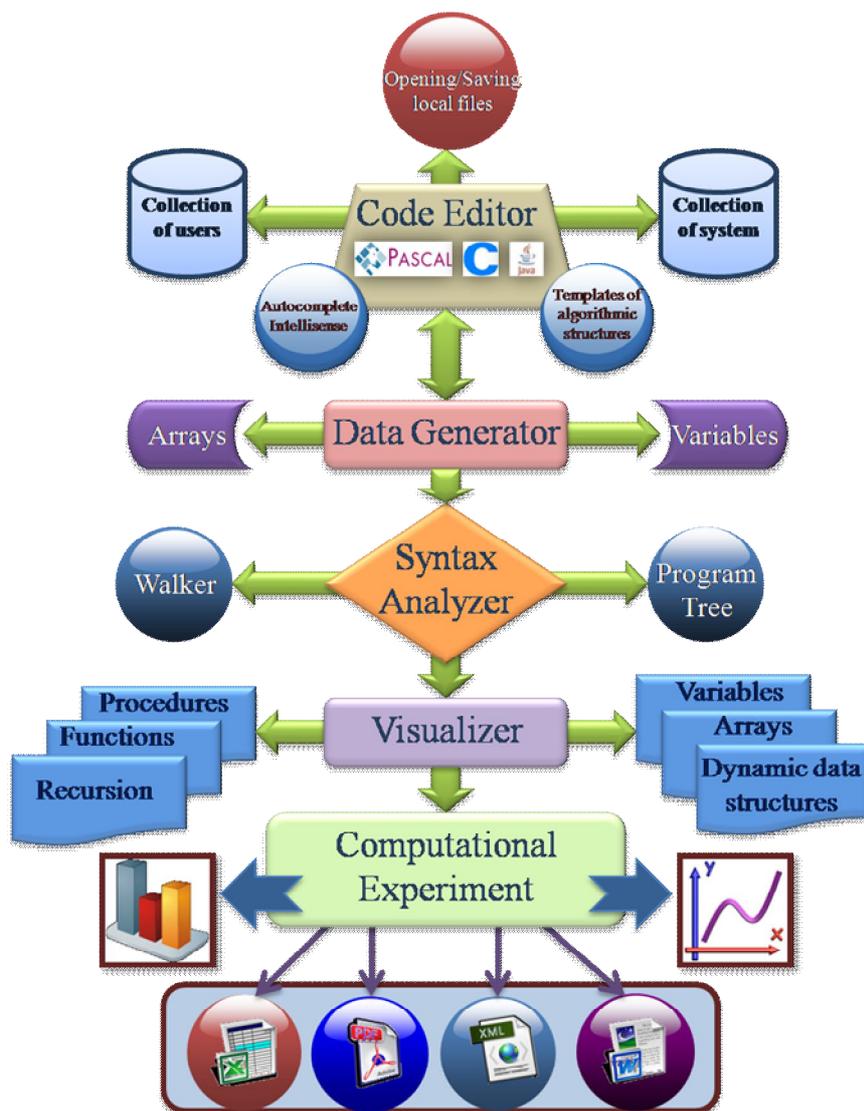


Рис.2. Схема взаимосвязи компонентов среды демонстрации программ

Вычислительный эксперимент по изучению эффективности алгоритмов проводится с помощью специального модуля «Среда демонстрации» интегрированной среды изучения курса «Основы алгоритмизации и программирования». В среде демонстрации пользователь имеет возможность выбрать и открыть алгоритм для демонстрации из коллекции системы или из коллекции пользователя, инициализировать данные для демонстрации, выполнить визуализацию работы алгоритма в непрерывном или пошаговом режимах.

Использование динамических образов операций присваивания, сравнения, передачи параметров в процедуры и функции, рекурсивных вызовов процедур и функций, процесса генерации входных данных, делает среду демонстрации исключительно полезным средством изучения основ алгоритмизации и программирования [8].

Под вычислительным экспериментом понимается метод изучения эффективности алгоритмов поиска и сортировки с помощью математического моделирования. Эксперимент предусматривает, что после построения математической модели проводится ее численное исследование, что позволяет воспроизвести поведение исследуемого объекта в различных условиях или в разных модификациях.

При аналитическом подходе сложность алгоритмов доказывается с помощью строгих точных выкладок. Однако на практике студенты очень тяжело воспринимают этот материал. Другой способ – в среде демонстрации выполнить каждый алгоритм на нескольких массивах данных и исследовать характеристики алгоритма (количество сравнений, количество перестановок и время выполнения), используя статистические возможности, графическую интерпретацию данных и возможности генерации отчетов модуля «Вычислительный эксперимент», и, таким образом, исследовать эффективность алгоритмов.

Вычислительный эксперимент позволяет студенту понять особенности алгоритмов поиска и сортировки и осознать зависимости, которые объясняют сложность алгоритмов.

Выводы

В результате проведенных исследований предложен подход концептуального моделирования информационных систем поддержки учебной деятельности, позволивший выполнить проектирование информационной системы для организации вычислительного эксперимента в задачах поиска и сортировки.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

- 2 Ermolayev, V., Keberle, N., Matzke, W.-E.: An Ontology of Environments, Events, and Happenings. In: 31st Annual International Computer Software and Applications Conference, pp. 539-546, IEEE Press, NY (2008).
- 3 Gruber T. R. A translation approach to portable ontology specifications // Knowledge Acquisition. – 1993. – № 3. – Vol. 6.
- 4 Guarino N., Poli R. The role of ontology in the information technology // International journal of Human-computer studies. Special issue on ontology, – 1995. – № 43(5/6). – P. 623-965.
- 5 Kleshchev A.S., Artemjeva I.L. «A structure of domain ontologies and their mathematical models» // In the Proceeding of The Pacific Asian Conference on Intelligent systems 2001 (PAIS 2001), Korea Intelligent Information Systems Society, 2001, pp. 410-420.
- 6 Uschold M., Gruninger M. ONTOLOGIES: Principles, Methods and Applications // Knowledge Engineering Review. –1996. – № 2. – 1996. – Vol. 11.
- 7 Web-среда для изучения основ алгоритмизации и программирования / Спиваковский А.В., Колесникова Н.В., Ткачук Н.И., Ткачук И.М. // Управляющие системы и машины. – 2008. – С. 70-75.
- 8 Основи алгоритмізації та програмування. Обчислювальний експеримент. Розв’язання проблем ефективності в алгоритмах пошуку та сортування: Навчальний посібник/А.В. Співаковський, Н.В. Осипова, М.С. Львов, К.В. Бакуменко. – Херсон: Айлант, 2010. – 100 с.: іл..
- 9 Співаковський О.В. Відеоінтерпретатор алгоритмів інтегрованого середовища вивчення курсу “Основи алгоритмізації та програмування” / Співаковський О.В., Колеснікова Н.В. // Нові інформаційні технології в освіті для всіх: система електронної освіти. – 2008. – № 3. – С. 399-404.
- 10 Шалфеева Е. А. Классификация структурных свойств онтологий // Искусственный интеллект. – 2005. – № 3. – С. 67-77.