

## Понятийный граф как основа ведения модели знаний

# 05, май 2011

автор: Гиря И. А.

УДК 004.9

Национальный исследовательский ядерный университет «МИФИ»  
[ivan.girya@gmail.com](mailto:ivan.girya@gmail.com)

### 1. Введение

С момента возникновения идеи об использовании компьютера в процессе обучения различные исследовательские группы занимаются вопросом повышения эффективности данного подхода. Несмотря на все достижения дисциплины HCI (*human-computer interaction*, человеко-машинное взаимодействие) остается множество нерешенных вопросов в тех аспектах процесса передачи знаний, которые отличают компьютерное обучение от «живого». Совокупность этих проблем можно объединить в единую задачу, требующую решения – задачу индивидуализации процесса компьютеризированного обучения.

В ходе классического обучения в формате «человек–человек» опытный преподаватель способен определить сильные и слабые стороны ученика, оценить его общую эрудированность и уровень знаний конкретной дисциплины – и, исходя из этих данных, определить объем и оптимальную форму подачи материала.

Обучение, основанное на использовании компьютерных систем, в силу своей специфики изначально было лишено возможности учитывать личные особенности каждого конкретного ученика, по сути представляя собой переложение бумажного учебника в цифровой вид с добавлением элементов интерактивности.

Решением этой проблемы заняты разработчики *адаптивных систем обучения*. Идея таких систем заключается в создании алгоритмов и инструментов накопления информации о пользователе (ученике), с целью использования этой информации при определении *учебной траектории* ученика по курсу. Кроме того, с учетом особенностей знаний и личностных характеристик ученика, могут регулироваться объем подаваемого учебного материала, а также форма его представления.

Важной подзадачей в проектировании адаптивной системы является определение концепции *модели ученика* – структуры, предназначенной для хранения информации о пользователе системы. Модель ученика является основным источником информации, на основе которой система осуществляет адаптацию учебного курса к пользователю.

Данная статья посвящена подходу к формированию модели знаний ученика, основанной на понятийном графе предметной области.

## 2. Структура модели знаний ученика

Важнейшей задачей модели ученика является описание его текущих знаний. Будем называть часть модели ученика, содержащую информацию об уровне его знаний *моделью знаний ученика*.

Поскольку невозможно определить представление знаний в компьютерной системе в том виде, который используется человеческим мозгом, разработчикам приходится обращаться к различным моделям представления знаний. Существуют различные способы определения структуры модели знаний [1].

Предлагается подход к ведению модели знаний ученика, основанный на модели предметной области. Такой подход представляется наиболее эффективным для задач обучения, так как является ориентированным на конкретное поле знаний обучающей среды, с которой придется контактировать пользователю.

Содержание любого учебного курса может быть представлено конечным набором *понятий*, соответствующих входящим в курс информационным единицам. Под понятием понимается словесный идентификатор, описывающий некий объект, явление, свойство или процесс, относящийся к теме курса [4]. Представим множество, соответствующее набору понятий, составляющих предметную область обучающей системы (т.е. программу учебного курса) в следующем виде:

$$M = I_1 \cup I_2 \cup \dots \cup I_m, m \in N$$

Здесь элементы  $I_i$  соответствуют информационным единицам (понятиям).

В этом случае множество, соответствующее знаниям ученика, является подмножеством выше приведенного «эталонного» набора знаний обучающей системы и определяется выражением:

$$M_s \subset M$$

Таким образом, задача разделяется на две:

- разработка эффективного механизма представления знаний предметной области (учебной программы курса);

- создание связи модели знаний ученика с моделью знаний предметной области.

Распространение концепции *lifelong education* (непрерывное обучение) и связанных с ней идей *lifelong user modeling* (непрерывное ведение модели пользователя) заставляет отойти от рассмотрения задачи в масштабе одиночного учебного курса [3]. Необходимо прийти к созданию такой структуры модели знаний, которая могла бы быть использована неограниченное число раз, постепенно накапливая и уточняя информацию об уровне знаний пользователя-ученика.

Понятийные графы являются одной из распространенных и удобных методик представления знаний предметной области обучающей системы [4].

Организация модели знаний адаптивной обучающей среды в виде масштабируемого понятийного графа представляется эффективной основой для моделирования пользовательских знаний, не ограниченного временем работы с конкретной обучающей системой.

На рис. 1. изображен расширяемый понятийный граф, представляющий модель знаний, объединяющей предметные области курсов А и В.

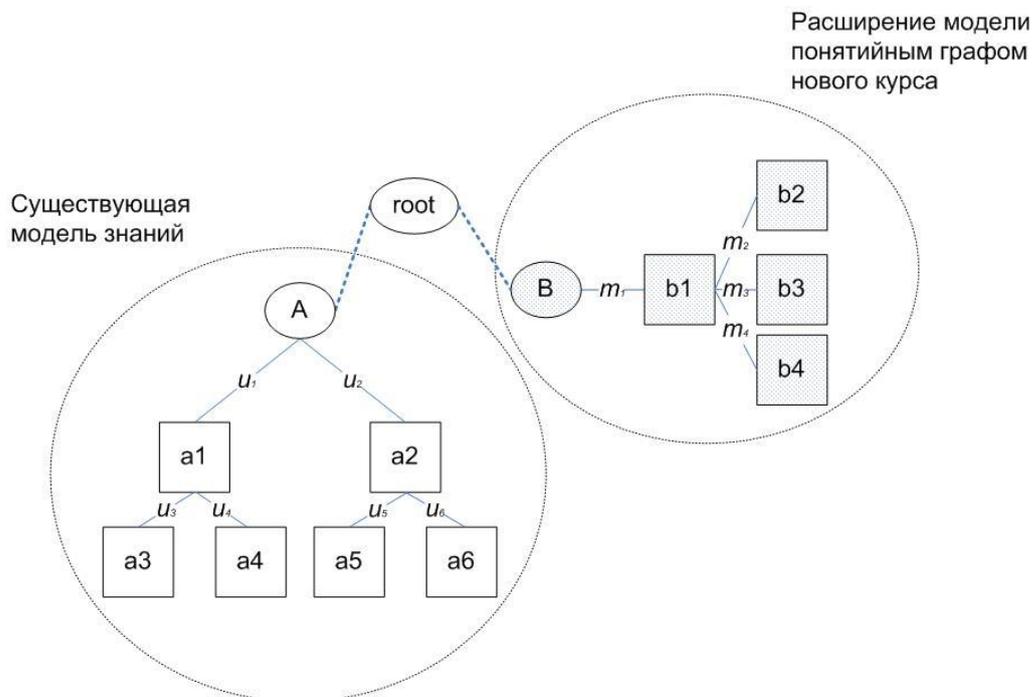


Рис. 1. Расширяемый понятийный граф

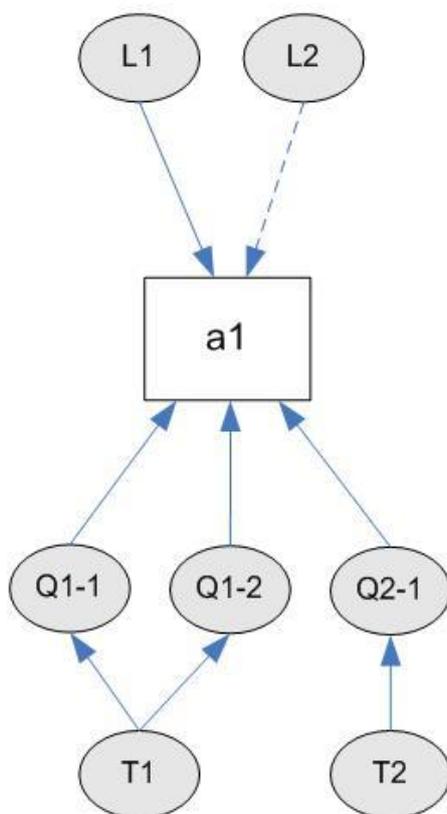
Элементы  $a1..a6$ ,  $b1..b4$  являются информационными единицами (понятиями) из курсов А и В. Линии  $u1...u6$ ,  $m1..m4$  соответствуют связям между узлами понятийного графа.

Элемент  $root$  является объединяющим понятием для курсов А и В. Например, для курсов «Термодинамика» и «Оптика» элемент  $root$  принял бы значение «Физика».

При появлении новых учебных курсов, существующий понятийный граф должен расширяться за счет добавления новых узлов, указания связей между ними и имеющимися узлами, а также, при необходимости, добавлении элементов типа  $root$ . Полученную структуру будем называть *деревом знаний* комплекса адаптивных обучающих курсов.

Выбранная структура модели знаний удобна возможностью настройки связей между узлами понятийного графа и различными объектами обучающих систем (такими как раздел, тема, веб-страница, тест, вопрос из теста и т.д.).

Пример описания таких связей для понятия  $a1$  приведен на рис. 2.



**Рис. 2. Связи понятия с объектами обучающей системы**

Здесь L – уроки учебного курса, в которых встречается понятие  $a_1$ . При этом сплошная стрелка, соединяющая урок с понятием, означает, что ученик, успешно изучивший данный урок, может считаться также и успешно освоившим понятие  $a_1$ . Прерывистая стрелка указывает на неполное освещение понятия в уроке.

T – тесты, состоящие из вопросов Q.

Таким образом, имея информацию об успешности продвижения ученика по учебному курсу (обсуждение технологий сбора такой информации лежит за рамками данной статьи), мы получаем возможность сделать выводы об успешности освоения учеником определенного набора понятий из дерева знаний. Так вычисляется подграф дерева знаний, соответствующий модели текущих знаний ученика.

В дальнейшем, по мере работы ученика с другими обучающими курсами, учебные программы которых в виде набора понятий включены в единое дерево знаний, модель знаний ученика будет расширяться параллельно с расширением самого дерева.

При этом, на основании характера связей между понятиями ( $u_1...u_b, m_1..m_4$  на рис.1.) есть возможность строить предположения об уровне знаний ученика не только о понятиях, непосредственно затронутых в учебном курсе, но и о связанных с ними понятиях.

### **3. Создание и применение понятийных графов**

Выделим два возможных способа создания понятийного графа обучающей системы:

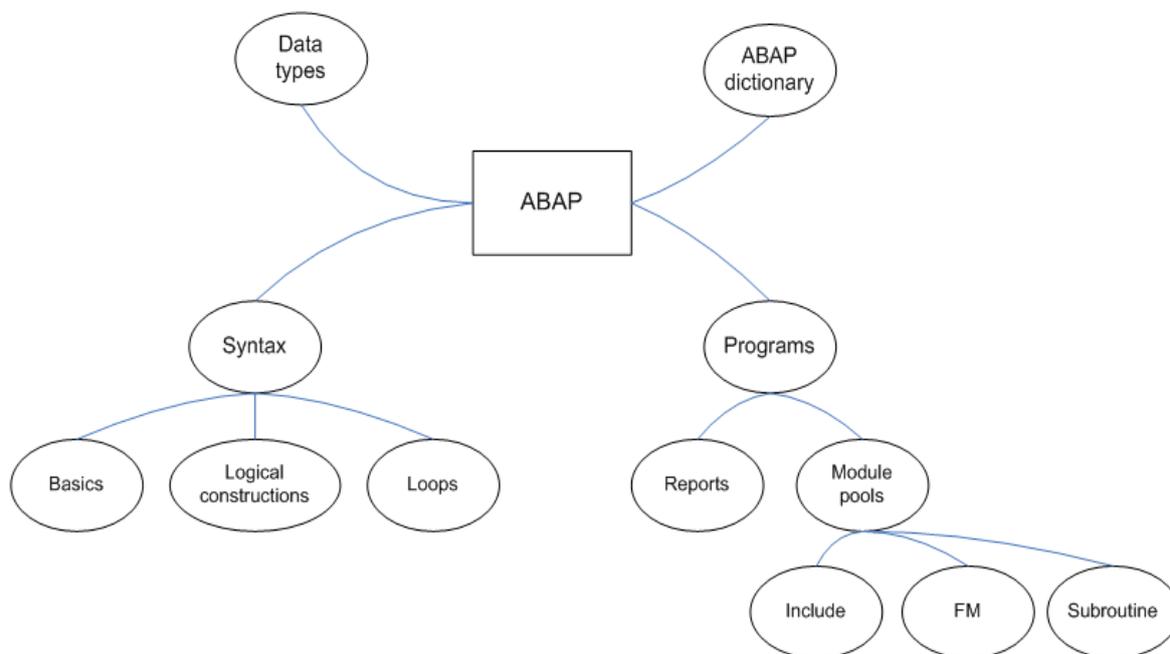
(1) использование автоматизированных систем формирования понятийной структуры исходных материалов;

(2) ручной ввод информации экспертом по предметной области (например, преподавателем курса).

При использовании способа (1) отрывок текста из учебного курса (например, лекция) обрабатывается методами text mining. Текст последовательно проходит стадии морфологического, статистического, синтаксического и семантического анализа. На выходе получается понятийная структура текста с указанием ключевых слов/словосочетаний и характеров отношений между ними. [5]

В случае использования способа (2), от эксперта требуется последовательно указать все понятия, формирующие модель знаний курса, определить тип отношений между ними, а также указать связи с существующими курсами и объектами курсов.

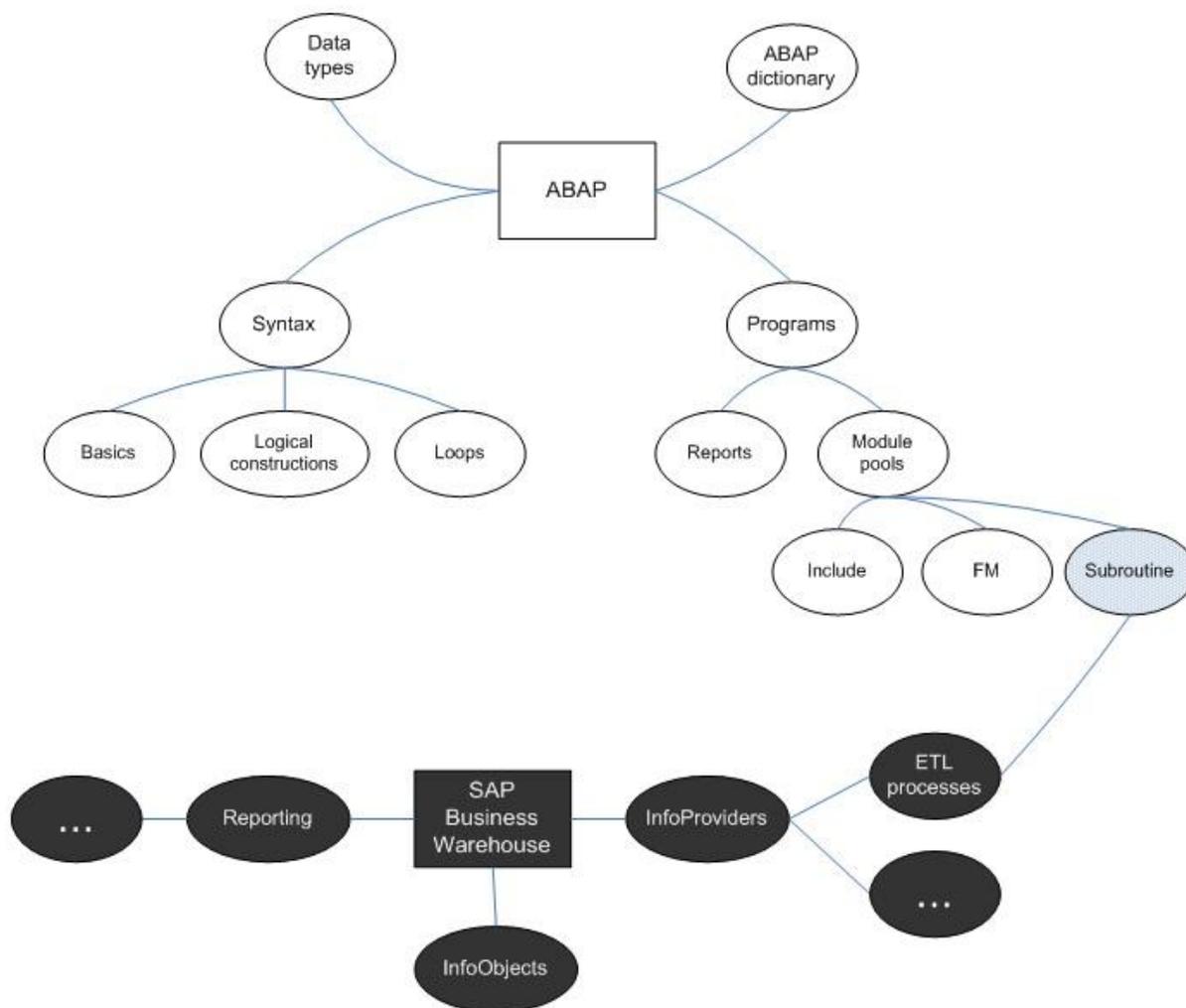
На рис.3. изображен понятийный граф, составленный по методу (2) для проведения тренинга «Основы ABAP» (*Advanced Business Applications Programming* – внутренний язык программирования продуктов компании SAP, в частности, SAP ERP).



**Рис. 3. Понятийный граф «Основы ABAP»**

При составлении данного графа мы руководствовались тем, что обучаемым (стажерам и младшим специалистам) в проектной работе по внедрению SAP более всего понадобятся практические навыки по написанию на ABAP функциональных модулей (подпрограмм) для обработки данных. Поэтому «ветви», посвященные синтаксису языка и конкретным типам программ развернуты более подробно.

Следующим шагом обучения специалистов был тренинг по SAP BW (*Business Warehouse* – решение SAP, сочетающее в себе хранилище данных и OLAP-технологии). Для его проведения был составлен новый понятийный граф. Упрощенный вид этого графа, а также его связь с понятийным графом «Основы ABAP» изображены на рис. 4.



**Рис. 4. Связь понятийных графов «Основы ABAP» и «SAP BW»**

Как видно из рисунка, роль узла «root» в данном случае отведена понятию «Subroutine», одинаково важному как в тренинге по ABAP, так и в курсе по BW.

Описанные выше графы предназначены для составления моделей знаний специалистов и их применения при проведении аттестации, основанной на методиках адаптивного тестирования.

Информация о знаниях хранится в виде кортежей следующего вида:

$$\langle S, T, L \rangle$$

Здесь S – имя ученика, T – понятие (узел в графе), L – показатель усвоения понятия (в процентах). Значение L определяется несколькими способами:

- 1) напрямую через тестирование – на основе успешности выполнения заданий, связанных с понятием T;
- 2) в случае успешного изучения всех понятий, стоящих ниже T по иерархии графа;

3) вручную преподавателем курса/руководителем сотрудника.

Ценность технических знаний специалиста внедрения SAP определяется возможностью их эффективного применения для нужд конкретного заказчика.

*Профессиональную компетентность* опишем кортежем:

$$K = \langle [A], [T], S, F \rangle$$

Здесь [A] – множество понятий, относящихся к компетенции, [T] – комплекс возможных задач и применений, S – бизнес-сценарий (например, анализ эффективности производственной линии), F – отрасль применения (энергетика, розничные продажи, банки и т.д.).

Элементы из [A] и [T] определяются конкретными узлами понятийного графа, связанными с рассматриваемой компетенцией. Вывод об определенной компетентности специалиста делался на основе проверки двух условий:

- наличие проектного опыта по данной компетенции (т.е. выполнение задач, основанных на владении компетенцией по крайней мере в роли младшего специалиста);

- количество успешно изученных узлов из [A] и [T] составляет не менее 75%.

#### **4. Заключение**

Целью описанной организации модели знаний обучаемого является централизованный сбор данных об уровне его подготовки. Принцип расширяемого понятийного графа позволяет избежать фрагментации сведений об ученике, рассредоточенных в различных локальных системах, что имеет определяющее значение при ориентации обучения на принципы *lifelong learning*.

Эта особенность предлагаемой структуры позволяет применять ее в построении компетентностных моделей знаний (КМЗ). В терминах базисной реализации КМЗ [6], узлы понятийного графа предстают *базисными понятиями* и *базисными методами* КМЗ.

В сочетании с производственными характеристиками (отрасль, бизнес-процесс и т.п.) информация об овладении специалистом базисными понятиями позволяет делать выводы о конкретных его компетентностях.

## Список литературы.

1. *Е.С. Белоус, В.А. Кудинов, М.Э. Желнин.* Современные модели представления знаний в обучающих системах. // Ученые записки. Электронный научный журнал Курского государственного университета. — 2010. Номер 1. С. 9—14.
2. *Б.Г. Киселев.* Модель поля знаний курса «Принятие управленческих решений». // Научная сессия МИФИ 2004. Сборник научных трудов. — 2004. Т.13204-205. — Б. ц.
3. *J. Key, B. Kummerfield.* Lifelong user modelling goals, issues and challenges. // Proceedings of the Lifelong User Modelling Workshop at UMAP'09 User Modeling Adaptation and Personalization. — 2010. p. 27—34.
4. *Г.А. Воробьев, П.А. Корнев, В.Н. Малыш.* Модель структуризации содержания обучения на основе построения иерархической понятийной сети знаний. // Вестник РУДН. — 2010. Выпуск 4.
5. *О.В. Окропишина.* Технология автоматизированного формирования понятийной структуры научного контента. // Электронные библиотеки: перспективные методы и технологии, электронные коллекции 2010. Сборник научных трудов конференции. — 2010.
6. *Ю.С. Васильев, В.Н. Козлов.* Компетентностные модели знаний и содержания высшего профессионального образования. // XIX Всероссийская научно-методическая конференция «Проектирование образовательных программ высшего профессионального образования на компетентностной основе». — 2009. Сборник 1.