

KNOWLEDGE BASE OF LEARNING SYSTEM CONSTRUCTION ON THE BASIS OF MODEL OF IMAGE THINKING

Oleg V. Bisikalo

Vinnitsya State Agricultural University

Abstract

The work is devoted to the creation of «clever» content for electronic textbooks. The principles of construction of image thesaurus are examined on the basis of requirements of the model of image thinking. The formal relation model of knowledge base of learning system, filled as a result of manipulation of texts with teaching material is offered. The operating prototype of the system is programmatically realized on the basis of technology of Python + SQLite.

ПОСТРОЕНИЕ БАЗЫ ЗНАНИЙ ОБУЧАЮЩЕЙ СИСТЕМЫ НА ОСНОВЕ МОДЕЛИ ОБРАЗНОГО МЫШЛЕНИЯ

Бисикало О.В.

Винницкий государственный аграрный университет

Аннотация

Работа посвящена созданию «умного» контента для электронных учебников. На основе требований модели образного мышления рассматриваются принципы построения базы знаний обучающей системы. Предлагается формальная реляционная модель словаря образов, наполняемого в результате обработки текстов с учебным материалом. Действующий прототип системы программно реализован в технологии Python + SQLite.

Введение в проблему. Лидирующая роль компьютерных технологий в научно-технической революции пока еще не изменила существующего положения, при котором значительная часть человеческих знаний заключена в текстовой форме. Вне зависимости от типа носителя информации символично-вербальный формат знаний остается стандартом для современной науки и образования [1]. Одной из проблем е-обучения является более быстрое ослабление внимания и, следовательно, пониженный уровень восприятия текстовой информации с экрана компьютера по сравнению с традиционным учебником. Улучшение качества мониторов пока еще не полностью решает данную проблему и, тем более, не учитывает возможностей компьютерного представления информации.

Одним из возможных путей улучшения восприятия электронного текста является моделирование образного мышления с целью представления отдельных слов в виде символических обозначений образов [2]. В таком случае простое повествовательное предложение или синтагма текста может рассматриваться как аналог события, где образы объединяются между собой с помощью ассоциативных связей [3]. Идея подхода заключается в том, чтобы дополнить традиционный гипертекст с парадигматическими связями возможностями синтагматического знания в виде дополнительных ассоциативных связей между образами. Поэтому **постановкой задачи** будем считать формализацию принципов построения глоссария как словаря образов предметной области исходя из требований общей модели образного мышления.

Требования к глоссарию с точки зрения модели образного мышления. Предлагаемый подход основан на моделировании естественного интеллекта путем самонаблюдения и объединения известных результатов таких наук, как физиология, психология, лингвистика, нейропсихология [2]. В концептуальную модель образного мышления заложены следующие постулаты, с которыми не должны входить в противоречие возможные технические решения словаря образов или глоссария предметной области.

В первую очередь необходимо отметить принцип самообучаемости живых носителей интеллекта [1]. Известно, что грудной ребенок практически с нуля обучается всей сложности мира по простому принципу, сформулированному еще в древнем Риме как «постепенность и постоянство». Для восприятия любого нового знания нужен хороший фундамент [4], поэтому словарь образов необходимо «вырастить» до требуемого уровня путем организационной опеки от наиболее общих понятий до специальных терминов. Иначе парадоксальной выглядит экспертная система, решающая сложные, но узкоспециализированные задачи, но бессильная перед «детскими» вопросами.

Существенным отличием естественной памяти от компьютерной является принципиальное отсутствие операции стирания информации [5]. Если человек встречает новый образ, то он никогда его не проигнорирует даже потому, что в отдельных случаях это чревато опасными последствиями. Точно также нельзя удалить уже имеющийся образ, например, негативный – случаи подсознательного вытеснения, как показал еще З. Фрейд, приводят

к патологиям в психике. Лечение может быть только «вытягивание» образа в сознание и «наращивание» на него положительной сети связей с другими образами. Данный принцип организации памяти можно, несколько упрощая, сформулировать, как «всех впускать – никого не выпускать».

Третий из основных постулатов модели касается механизма пополнения словаря образов. Известно, что при восприятии нового образа у человека на подсознательном уровне включается ориентировочный рефлекс [6]. Одной из задач этого рефлекса является построение сети связей нового образа с уже существующими в памяти образами. Другими словами, происходят процессы распознавания или классификации образа. Любой новый образ принципиально не может быть независимым, поскольку в момент ввода в систему «обрастает» максимально возможным количеством связей с другими образами. Тем самым проявляется рекуррентная сущность определения образа – «скажи мне, кто твои друзья, и я скажу, кто ты».

Принципы построения словаря образов. С учетом рассмотренных требований модели образного мышления предлагаются следующие принципы создания глоссария образов на основе внесения в систему текстов учебного содержания.

1. Обучающие тексты (учебные дозы) и составляющие их фразы (синтагмы) представлены в системе как отношения

$$\text{Text} - RE \subset \text{Text} - Id \times Bi - Te \times Title \times Author \times Time, \quad (1)$$

где *Text-Id* – уникальный код текста, *Bi-Te* – двоичный код учебной дозы, *Title* – название текста, *Author* – автор текста, *Time* – время внесения дозы в систему и

$$\text{Event} \subset \text{Event} - Id \times Bi - Sy \times \text{Text} - Id \times \text{Syntagma}, \quad (2)$$

где *Event-Id* – уникальный код синтагмы, *Bi-Sy* – двоичный код учебной фразы, *Syntagma* – вербальное обозначение фразы.

2. Собственно словарь образов представлен в виде отношения

$$\text{Image} \subset Bi - I \times \text{Object} - Quality \times \text{Object} \times \text{Notion} \times \text{Method} \times \text{Method} - Quality, \quad (3)$$

где *Bi-I* – двоичный код образа и вербальные обозначения: *Object-Quality* – качество объекта, *Object* – объект, *Notion* – понятие, *Method* – метод, *Method-Quality* – качество метода.

3. Элементарным конструктом предложения является ассоциативная пара образов, представленная в виде отношения

$$Assoc - Twice \subset Bi - I_1 \times Bi - I_2 \times Twice - Id \times Force \times Force^{-}, (4)$$

где $Bi - I_1$ – двоичный код первого образа пары, $Bi - I_2$ – двоичный код второго образа пары, $Twice - Id$ – уникальный код пары, $Force$ – значение силы прямой связи между образами, $Force^{-}$ – значение силы обратной связи между образами.

4. Ввод в систему обучающей информации осуществляется путем внесения данных о парах в такие отношения, как тип связи
- $$Link \subset Link - Id \times Link - Type \times Specification, (5)$$

где $Link - Id$ – уникальный код типа связи, $Link - Type$ – вербальное обозначение типа связи, $Specification$ – правила применения типа связи и вопросительное местоимение

$$Inter - Pronoun \subset Pronoun - Id \times Link - Id \times Pronoun, (6)$$

где $Pronoun - Id$ – уникальный код местоимения, $Pronoun$ – вербальное обозначение местоимения.

5. Вопросительное местоимение между образами пары задается двумя способами. Сначала можно выбрать тип связи из 7 кортежей отношения $Link$ (определение, сказуемое, подлежащее, обстоятельство места, обстоятельство времени, обстоятельство, дополнение), тогда выбранный тип служит фильтром, и количество возможных местоимений $Inter - Pronoun$ уменьшается. С обратной стороны, если выбран сначала вопрос как кортеж из $Inter - Pronoun$, то для контроля пользователю автоматически демонстрируется соответствующий ему тип связи из $Link$.
6. Пользователю предоставляется возможность выбрать каждое слово пары в меню, составленного из слов текущего предложения $Event$, а затем указать роль соответствующего образа в синтагме (качество объекта, объект, понятие, метод, качество метода). Слова представлены в виде отношения
- $$Words \subset Word - Id \times Word \times Bi - I, (7)$$
- где $Word - Id$ – уникальный код слова, $Word$ – собственно слово, $Role - Id$ – уникальный код роли слова в синтагме, а роли в виде отношения
- $$Role \subset Role - Id \times Role - Type, (8)$$
- где $Role - Type$ – вербальное обозначение роли.
7. С целью привязки выбранного слова пары к образу составляется ранжированный список наиболее похожих внешне слов из уже существующих в словаре образов $Image$. При составлении

списка [5] учитываются правила синтаксиса, например если идет речь о сказуемом, то в список нужно помещать атрибуты *Method* или *Notion*. Пользователю предоставляется возможность выбрать в меню опцию, при необходимости откорректировать соответствующую статью словаря образов или ввести совершенно новую.

Следует также отметить, что самостоятельную ценность в предложенной реляционной модели имеют отношения *Image* и *Assoc-Twice*, поскольку они совместно образуют ассоциативную сеть образов и несут главную смысловую нагрузку. Все остальные отношения фактически моделируют ленту событий или долговременную память [3].

Программная реализация тезауруса образов. С целью обеспечения многопользовательского доступа к тезаурусу образов предлагается содержание в одном файле двух таблиц *Image* и *Assoc-Twice* как ядра системы, отдельное его администрирование и организация к нему параллельного доступа on-Line. Окончательное формирование базы данных тезауруса достигается с помощью отношения *Construct*, в кортежах которого фиксируются особенности использования одинаковых ассоциативных пар образов в различных предложениях:

$$\text{Construct} \subset \text{Construct} - Id \times \text{Pronoun} - Id \times \text{Twice} - Id \times \text{Word} - Id_1 \times \text{Word} - Id_2 \times \text{Event} - Id \times \text{Role} - Id, \quad (9)$$

где *Construct-Id* – уникальный код конструкта синтагмы. При создании новой записи в таблице *Construct* параллельно в соответствующую запись таблицы *Assoc-Twice* обязательно добавляется 1 или в поле *Force* или в поле *Force'* в зависимости от того, в прямом или в обратном порядке сработала данная ассоциативная связь.

Действующий прототип тезауруса образов программно реализован на основе технологии Python + SQLite, которая соединяет возможности языка запросов SQL к реляционной базе данных с парадигмами объектно-ориентированного и функционального программирования. На рис. 1 представлена схема данных тезауруса образов, используемая в SQLite.

Рассмотренная схема данных соответствует реляционной модели (1)÷(9) тезауруса образов и отвечает требованиям концептуальной модели образного мышления [2].

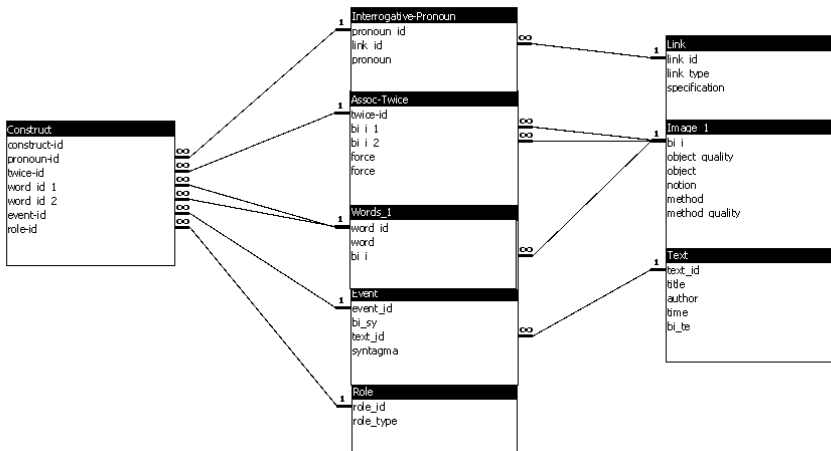


Рис.1. Схема данных тезауруса образов.

Тестовый пример предложенного подхода реализован на основе внесения в обучающую систему известного текста А. С. Пушкина «Сказка о рыбаке и золотой рыбке». На рис.2 представлен внешний вид оболочки обучающей системы, заполненный тестовыми синтагмами.

↵ AOL Radio - Featuring 200+ Free ... <http://92.112.44.69:8080/>

| [events](#) | [images](#) |



они жили веткой землянке ровно тридцать лет три года X

старик ловил неводом рыбу X

старуха пряла свою пряжу X

раз он море закинул невод X

пришел невод одною тиною X

он другой раз закинул невод X

пришел невод травкою морскою X

третий раз закинул он невод X

пришел невод одною рыбкой непростюю золотую X

буря мглою небо кроет X

жил старик своею старухой самого синего моря X

как взмолился золотая рыбка голосом молвит человечим X

отпусти ты старче меня море X

дорогой себя дам откуп X

откуплюсь чем только пожелаешь X

удивился старик испугался X

Рис.2. Внешний вид оболочки обучающей системы.

На рис.3 и рис.4 представлены режим ввода синтагмы в оболочку и режим графической интерпретации связей синтагмы.

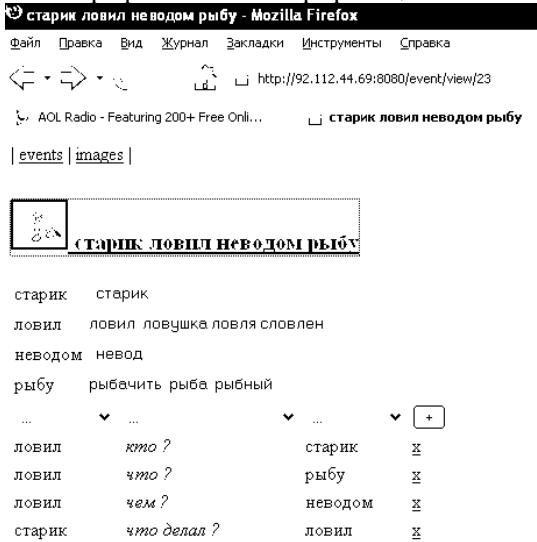


Рис.3. Режим ввода синтагмы в оболочку обучающей системы.

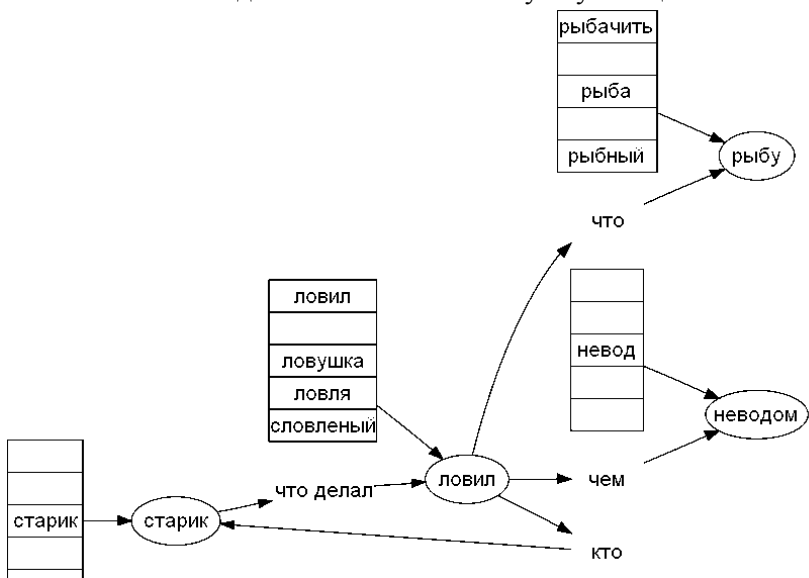


Рис.4. Режим графической интерпретации связей синтагмы.

В таблице 1 показаны наиболее близкие синтагматические связи образа «невод», полученные в результате запроса в SQLite:
 select total(force) as f, * from
 (select force, b1 as b from a_t where b2=27
 union all
 select force, b2 as b from a_t where b1=27), image
 where bi_i=b group by b order by f desc.

Табл.1. Образы-методы, ассоциативно близкие к образу «невод».

3	35	35	пришедший	{ null }	{ null }	прийти	{ null }
3	33	33	кинутый	{ null }	{ null }	кидать	закинуто
1	26	26	словленный ловушка ловля			ловить	{ null }

Выводы. В работе намечены основы подхода к созданию «умного» контента для электронных учебников. На основе требований концептуальной модели образного мышления рассмотрены принципы построения словаря образов, дополняющего обычный гипертекст ассоциативными связями синтагматического типа. Предлагается формальная реляционная модель глоссария, наполняемого в результате обработки текстов с обучающим материалом. Для программной реализации обучающей системы выбрана технология Python + SQLite.

Следует отметить, что новые возможности предложенного подхода [7,8,9,10] достигаются за счет большей трудоемкости внесения текстов учебного содержания в систему. Однако сравнительно небольшие объемы обучающего контента [10] не позволяют считать такое ограничение критическим.

Перспективным направлением дальнейших исследований может быть усиление максимальной схожести и возможностей корректирования образа в момент появления новой информации. При невозможности в спорных случаях привязать слово к образу в список для выбора должны попадать те кортежи отношения *Image*, которые наиболее близко связаны с уже распознанным парным и другими образами предложения как инсайтные.

Литература

1. Манак А.Ф., Манак В.В. Електронне навчання і навчальні об'єкти. – К.: ПП "Кажан плюс", 2003. – 334 с.
2. Бисикало О.В. Принципы построения концептуальной модели образного мышления. Первая международная конференция «Новые информационные технологии в образовании для всех»,

расширенные материалы конференции (29-31 мая 2006 г.). – Киев, 2006. – с. 25-34.

3. Бисикало О.В. Принципы построения лингвистической системы на основе модели образного мышления. В сборнике «Труды восьмой международной научно-практической конференции «Современные информационные и электронные технологии», г. Одесса, 21-24 мая 2007 года». – Одесса, СИЭТ–2007, 2007. с.63.
4. Бисикало О.В. Подход к построению ассоциативной памяти на основе модели образного мышления. В материалах Международной научно-технической конференции «Искусственный интеллект. Интеллектуальные системы (ИИ – 2007)» 24-29 сентября 2007 г. – Донецк-Таганрог-Минск, 2007. – с. 6-10.
5. Бісікало О.В. Структура блоку пам'яті на основі моделі образного мислення людини // Искусственный интеллект – 2007. – № 3. – с. 461-468.
6. Бисикало О.В. Подход к моделированию образного механизма оперативной памяти. В сборнике трудов Второй международной конференции «Новые информационные технологии в образовании для всех: состояние и перспективы развития» 21-23 ноября 2007 г». – Киев, 2007. – с. 336-344.
7. Бісікало О.В. Алгоритм пошуку інформації на основі моделі асоціативної пам'яті людини. В збірнику «ІНТЕРНЕТ–ОСВІТА–НАУКА–2006, п'ята міжнародна конференція ІОН–2006, 10–14 жовтня, 2006. Збірник матеріалів конференції. Том 2.» – Вінниця, УНІВЕРСУМ-Вінниця, 2006. с.560-565.
8. Бісікало О.В. Система «питання–відповідь» в межах моделі образного мислення. В збірнику «Матеріали XIII Міжнародної конференції з автоматичного управління (Автоматика-2006)», м. Вінниця, 25-28 вересня 2006 року. – Вінниця, УНІВЕРСУМ-Вінниця, 2007. с.537-542.
9. Бісікало О.В. Класифікація образного пошуку. В збірнику «Тези доповідей Першої міжнародної науково-технічної конференції «Інтеллектуальні системи в промисловості і освіті – 2007» 7-9 листопада 2007 року». – Суми, 2007. с.14-15.
10. Бісікало О.В. Проектування електронного підручника на основі формалізації пізнавальної діяльності людини. В збірнику наукових праць «Перспективні технології навчання та освітні простори», вип. 1. – Київ, МННЦ ІТiС, 2007. – с. 179-190.