

Жаров Валентин Константинович

д-р пед. наук, профессор, заведующий кафедрой

Крылов Кирилл Анатольевич

студент

Институт информационных наук

и технологий безопасности

ФГБОУ ВПО «Российский государственный

гуманитарный университет»

г. Москва

**АВТОМАТИЗИРОВАННАЯ ОБРАБОТКА УЧЕБНЫХ ТЕКСТОВ
В ЭЛЕКТРОННЫХ СРЕДАХ И ПОВЫШЕНИИ ЭФФЕКТИВНОСТИ
ИХ ВОСПРИЯТИЯ В ПРОЦЕССЕ ОБУЧЕНИЯ: МЕТОДИКО-
МАТЕМАТИЧЕСКИЙ АСПЕКТ**

***Аннотация:** в данной статье авторами представлен алгоритм решения задачи выделения текстового ядра и получения методики формальной оценки соответствия учебного текста к тезаурусу предметной области, основанный на методе латентно-семантического анализа.*

***Ключевые слова:** электронно-образовательная среда (ЭОС), организация учебного процесса, структура информации, моделирование обучающей среды, оптимизация учебного процесса, ЛСА, семантические поля, атомарные понятия, атомарные словари.*

ЭВМ играют в настоящее время одну из ключевых ролей в моделировании процессов познания. Одной из задач попыток провести содержательные образовательные реформы последних лет, является повышение качества и объективности проверки знаний учащихся. Поэтому повышение эффективности образовательной среды вуза по замыслам авторов реформ, лучше – преобразований в образовании, связано с информационно-административной лабильностью и структурированностью учебной и организационной информацией [1]. Естественен во-

прос, а хорошо ли если коммуникационные технологии заменяют непосредственное общение между Учителем и Учеником? А хорошо ли переносить на «плечи» дистанционного обучения, например, на онлайн или скайповое обсуждение решений контрольных работ? Возможно ли точное представление о «возможностях технологий» в ЭОС (электронно-образовательной среде) в промежутках между «возбуждениями» управленческих структур министерства образования (Тому пример, Стандартизация всего образования от версии 2+ до, как теперь говорят, версии 3++. Или было бы интересно выяснить, а может быть компетентность без профессионального или иного образования? А компетентность – это чиновничье восприятие знаний? Достаточно спросить у чиновников от образования: а может быть глубокая или совершенная компетентность? Владение же знанием может быть и глубоким, и совершенным и др.!? Иными словами, где проходит граница между автоматизированным и автоматическим в образовательном процессе? В поисках ответов на поставленные вопросы в частности, проблемы автоматизации различных разделов (частей) учебного процесса обсуждаются широко и исследуются по различным направлениям. Данная статья посвящена и является продолжением нескольких статей авторов в данном дискуссионном журнале [2–4].

Одним из важнейших вопросов образовательного процесса всегда являлось построение учебного плана, по которому студенты смогут осваивать дисциплину. На сегодняшний день не существует общепринятого алгоритма автоматизации решения данной проблемы, способного давать объективную оценку эффективности изучения материала, следуя учебному плану. Очевидно, что проблема кроется в недостаточной проработанности модели обучения профессиональному восприятию и достаточной обработки «получаемой» учебной информации.

Рабочая программа дисциплины (РПД по старому УМК) представляет собой комплексную и идеальную модель изучения некоторой предметной области входящую в учебный план направления, в нашем случае – математики (направление прикладная математика). Любой, кто писал РПД, кто преподавал, знают, что в

него входят: список тем, освещаемых преподавателем либо по предпочтениям, либо по необходимости; в рамках же связанного учебного плана фиксируется набор учебных дисциплин, последовательность изучаемых, формы аттестации по ним, рубежному контролю знаний, количеству учебных часов и т. д. Получается Учебный план – сложная система, наши же исследования направлены на поиски общих (инвариантных) элементов, позволяющих ускорять познавательный процесс и повышать эффективности образовательной среды вуза, но не полное отрицание сложившейся образовательной системы, но достаточное его переосмысление ведущего к достижению целей и решение вышеопределенных задач.

Прежде всего, нужно сделать ряд уточнений: 1) одной из главных составляющих эффективного учебного плана являются тексты, они и стали объектом нашего исследования; 2) рассматриваемые тексты будут представлены в виде семантических моделей (с помощью семантических полей), в форме, пригодной для обработке на ЭВМ; 3) семантическая модель текста нами будет пониматься, как математическая модель, отражающая смысловые связи множеств слов, образующих тему; 4) нами рассматриваются только учебные тексты, используемые при подготовке студентов по математическим дисциплинам.

Разумеется, от Ученика потребуются стремление к эффективному обучению (В этой статье мы не обсуждаем проблему мотивации к обучению, считая студента идеальным.), то есть, он должен стремиться постигать читаемый ему предмет, как на занятиях в вузе, так и самостоятельно расширять свои знания в преподаваемой ему области. Разумеется, за соблюдением этих условий тоже важно проследить, но они лежат для нас в других областях исследования, в частности, психологии, поэтому в данной статье не рассматриваются «побочные» вопросы обучения. А вот на эффективное постижение логической структуры дисциплины, её терминологии и методологии качество образовательного процесса влияет напрямую, поэтому ответственность за обеспечение должного уровня обучения ложится на «плечи» организации учебного процесса.

Если представить образовательный процесс иерархически, то становится очевидным, что, как уже было сказано выше, подбор источников литературы по

изучаемой дисциплине играет одну из ключевых ролей. Это позволяет нам создать надлежащую терминологическую базу и сформировать у Ученика, Учителем конкретное понимание того, какие должны быть получены результаты по окончанию обучения, т.е. какими знаниями навыками должен обладать студент, прослушавший курс. Это важно, поскольку студенты редко могут самостоятельно определить ключевые элементы изучаемого предмета и порядок их изучения, но преподаватели с небольшим педагогическим опытом (конечно владеющих конкретной учебной дисциплиной), могут допустить недооценку возникшей образовательной среды, особенно, если дело касается первокурсников.

Процесс оценки эффективности учебного текста весьма трудоемкий и длительный, поскольку текст необходимо внимательно прочесть, что требует большого количества времени. Разумеется, по каждой дисциплине существует большое количество хорошо зарекомендовавшей себя, известной в ученых кругах литературы, и тем не менее, прогресс не стоит на месте, пишутся новые книги, переиздаются классические и для преподавателя это серьёзный повод задуматься о включении в свою учебную программу новой литературы, на замену или в дополнение к уже имеющейся. Оценка результатов решения по модернизации списка литературы потребует очень много времени, которого в современном сверхскоростном потоке информации нет, а значит, даже частичная автоматизация (т.е. оценки небольшого количества литературы) позволит существенно уменьшить его и получать, если не окончательную, то более объективную конкретную оценку принятого решения.

Исследование семантической модели текста строится с помощью выделения минимальных словарей и с последующей проверки их на соответствие предметному тезаурусу.

Основой алгоритма послужил метод латентно-семантического анализа данных (далее – метод ЛСА), который был нами адаптирован для случая выделения атомарного словаря из семантической структуры текста. В свою очередь, метод ЛСА основан на теореме Дж. Форсайта о сингулярном разложении матрицы, которая, в частности, на практике позволяет найти ядро матрицы вне зависимости

от её ранга. В нашем случае, последнее позволяет выделять атомарный словарь из учебного текста произвольного объёма, который является также текстовым ядром, соответствующим числу линейно-независимых строк матрицы [5].

Как готовится текст? Из рассматриваемого нами текста удаляются все слова не являющимися существительными, далее отбрасываются имена собственные, оставшиеся существительные сравниваются с известным профессиональным тезаурусом. После чего результат оценивается с помощью семантического словаря [6; 7]. Получившееся же множество понятий знаний из данного текста, понятно, что оно (множество) определяет более общее знание к исходному тексту, но целью всех преобразований текста был атомарный словарь, данного конкретного текста. Далее рассматриваются следующие тексты из той же учебной или научной дисциплины, например, учебной дисциплины общая алгебра, опять выделяется атомарный словарь нового текста и эта процедура повторяется с другими текстами сколько необходимо раз. Получается, таким образом, множество «атомарных» словарей данного множества текстов дисциплины. Но каждый из полученных атомарных словарей сам является текстом. Объединив тексты атомарных словарей с учетом пересечения понятий, повторим алгоритм поиска атомарного словаря для нового только что полученного текста. Таким образом, в результате применения описанного алгоритма получим атомарный словарь знания изучаемой дисциплины.

Рассматривая все атомарные словари математического блока дисциплин мы выделяем атомарный словарь данного блока. Далее мы согласуем волевым усилием, например, заведующего кафедрой или декана понятийный аппарат, имеющий специфические расширения в каждой математической дисциплине, т.е. осуществляем обратный процесс расширения тезауруса математического словаря, но в едином семантическом пространстве и тем самым унифицируем восприятие знаний по соответствующим разделам, например математики.

Каждый из полученных атомарных словарей на первом этапе (сужение до общего атомарного словаря) составлен, согласно алгоритму, заранее, опираясь

на существующие семантические словари дисциплины. Но, тогда, двигаясь в обратную сторону, можно будет наглядно построить уровни тезауруса во взаимосвязях понятий, например, так сделано в [8].

Каждое слово (гнездо) полученного уровня имеет соответствующую этому уровню степень абстракции. Очевидно, что возникновение (вывод) понятий уровня согласуется с правилами вывода принятых в данном научном знании. Далее строя иерархию понятий и определяя связи между понятиями, в зависимости от целей и задач обучения (бакалавриат, магистратура или специалитет) определяем стратегию и методическую основу обучения. Поскольку электронная система выделения атомарных словарей на первом этапе эксперимента (два года назад) была автоматизированной, то следующим шагом в эксперименте (этот учебный год) была автоматизация построения атомарных словарей и сопряжение этой системы с ЭОС института (ИИНТБ ФИСБ РГГУ).

В конечном счете, такой подход унификации знаний различных блоков учебных дисциплин позволят нам создать «базовый» атомарный словарь знаний первого курса по направлению прикладная математика, в который будут входить отчасти и слова гуманитарного цикла. Причем продолжая от курса методики расширения знания, мы памятуя, на этапе «подключения» к учебному плану например Философии можем с некоторыми допущениями и создать её атомарный словарь. Например, так «У философии вообще нет своего языка. Язык есть у философствующего. В этом я вижу одно из главных отличий философии от науки. Метаязык философии – такая же популярная фикция, как, скажем, «язык описания мира» [9, с. 21].

Используя прием связи информационно-педагогическими средами можно представить стратегию развития обучения от ступени бакалавров к ступени магистров. Или выделить блок профессиональной терминологии, например, удовлетворяющего тезаурусу защиты информации и «вводить» профессиональный лексикон студенту на первых курсах обучения причем так чтобы соблюдалась цельная картина будущей специальности. Таким образом, становится более ме-

нее ясным прохождением плавающей границы между автоматизированным и автоматическим ЭОС в процессе обучения профессиональному знанию. Понятно, что эмпирическое знание приобретаемое на последних курсах обучения будет строгим экзаменатором теоретического, что дает в методический арсенал высшей школы еще одну методику адаптации практического к теоретическому знанию

Практика применения метода ЛСА уже имела место в США, где был изобретен метод, однако в нашей стране такой практики до недавнего времени не было, хотя первые теоретические разработки в этом направлении были сделаны еще в 1968 году. В последние несколько лет на базе ИИНТБ ФИСБ РГГУ и ВШЭ ведутся исследования по созданию ЭОС с автоматическим анализом различных блоков дисциплин в сочетании с таким же управлением потоком информации в систему Деканат.

Формируя поле текстов, доступных студенту, можно добиться существенного повышения эффективности учебного курса, поскольку становится возможным управлять изменениями в лексиконе студента, соотносясь с профессиональным тезаурусом. Упростив же с помощью общего атомарного словаря взаимодействие между различными учебными дисциплинами в процессе обучения с учетом учебного плана и соответствующими РПД мы добиваемся эффективного терминологического единства, например, в преподавании математики по различным направлениям.

Список литературы

1. Информационно-педагогическая среда современного вуза: Коллективная монография / Под общ. Ред. В.К. Жарова. – М.: Янус-К, 2011. – 266 с.
2. Жаров В.К. О некоторых проблемах образовательной среды современного вуза // Образовательная среда сегодня: стратегия развития: Материалы IV Междунар. науч.-практ. конф. (Чебоксары, 08 дек. 2015 г.). – Чебоксары: ЦНС «Интерактив плюс», 2015. – №3 (4). – С. 183–186.
3. Жаров В.К. Об атомарных понятиях, методиках обучения и организации учебного процесса / В.К. Жаров, К.А. Крылов // Педагогический опыт: теория,

методика, практика: Материалы V Междунар. науч.-практ. конф. (Чебоксары, 25 дек. 2015 г.). – Чебоксары: ЦНС «Интерактив плюс», 2015. – №4 (5). – С. 164–167.

4. Жаров В.К. Применение математического структурирования в оптимизации учебных планов / В.К. Жаров, К.А. Крылов // Инновационные тенденции развития системы образования: Материалы V Междунар. науч.-практ. конф. (Чебоксары, 5 февр. 2016 г.). – Чебоксары: ЦНС «Интерактив плюс», 2016. – С. 88–90.

5. Форсайт Дж. // Численное решение систем линейных алгебраических уравнений / Форсайт Дж., Молер К. – М.: Мир, 1969 – 167 с.

6. Шведова Н.Ю. Русский семантический словарь. – М.: Азбуковник, 1998.

7. Караулов Ю.Н. Русский семантический словарь: Опыт автомат. построения тезауруса: от понятия к слову / Ю.Н. Караулов, В.И. Молчанов, В.А. Афанасьев, Н.В. Михалев. – М.: Наука, 1983. – 566 с.

8. Жаров В.К. Курс лекций по математике / В.К. Жаров, О.А. Матвеев, А.С. Панкратов, А.А. Роганов. – Вып. 1. – М.: Януск-К, 2008. – 110 с.

9. Жаров В.К. Курс лекций по математике / В.К. Жаров, О.А. Матвеев, А.С. Панкратов, А.А. Роганов. – Вып. 2. – М.: Януск-К, 2008. – 142 с.

10. Жаров В.К. Курс лекций по математике / В.К. Жаров, О.А. Матвеев, А.С. Панкратов, А.А. Роганов. – Вып. 3. – М.: Януск-К, 2008. – 132 с.

11. Пятигорский А.М. Мышление и наблюдение. Четыре лекции по обсервационной философии. – СПб.: Азбука, 2016. – 190 с.