

Экспертные системы идентификации текста для педагогических программных средств

Образовательные информационные технологии становятся все более заметным фактором развития педагогической науки и практики. Успешность современных инновационных процессов зависит не только от профессионально-педагогического мастерства преподавателей, состояния учебно-материальной базы образовательных учреждений, но и от уровня интеллектуализации информационных технологий. Стали значительно «умнее» для использования в учебном процессе операционные системы, сетевые технологии, офисные пакеты. Современные Web-технологии, мультимедийные технологии презентаций, Flash-анимации позволяют самостоятельно создавать электронные учебно-справочные пособия и демонстрационные материалы. На новый уровень выходят и педагогические программные средства (ППС) – средства создаваемые программированием на любом из объектно-ориентированных языков с использованием все тех же гипертекстовых и мультимедийных технологий. От материалов, самостоятельно создаваемых преподавателем, их отличает более сложная логика взаимодействия с обучаемыми, более широкий спектр педагогического инструментария.

В их развитии можно увидеть два направления: 1. ППС с фиксированным содержанием и управлением (защищены от изменений). 2. ППС с возможностью изменения содержания и управления (наличие инструментальных средств). В психолого-педагогической литературе оба направления подвергаются жесткой и не обоснованной критике. Считается, что диалог с массивом формализованной информации принципиально быть не может¹. Машинный контроль не выявляет способа получения результата, затруднений, типичных ошибок и других нюансов, которые не проходят мимо внимания педагога при устном и письменном контроле². Негативные оценки вызывают ППС, основывающиеся на примитивно-алгоритмической парадигме процесса обучения. Поэтому в педагогической науке усиливается потребность в исследованиях экспертно-эвристических ППС.

ППС первого направления, как правило, создаются авторскими коллективами из различных специалистов, обеспечивающих их высокий уровень исполнения. Не смотря на использование в разработке таких ППС современных достижений науки и технологий, процесс их создания часто занимает много времени и не может гибко реагировать на быстро изменяющейся потребности практики обучения. Учитывая постоянное изменение содержания многих учебных дисциплин, невозможность предварительного учета уровня и состояния подготовленности конкретных групп обучаемых на сегодняшний день более актуальными становятся задачи исследования ППС второго направления. Педаго-

¹ Харламов И.Ф. Педагогика. –М.: Гардарики, 2003. –С.270.

² Сластенин В.А. и др. Педагогика. –М.: Изд.центр Академия, 2003. –С.280.

гическое и программное развитие таких ППС должно отвечать профессиональным потребностям преподавателей, эффективно выражать их профессиональный опыт, предоставлять удобные инструменты для автоматизации обработки учебной информации, управления процессом обучения. Среди множества аспектов этого направления наибольший вклад на сегодняшний день могут внести экспертные системы идентификации текста. Это основывается на обязательности хранения и использования знаний в виде текста в общей системе аудиовизуальной и другой учебной информации.

Одним из примеров разработки экспертных систем для обучения является, разработанный автором статьи, язык идентификации текстовых структур в инструментальной системе тестирования ЭЛИС. В этой системе идентификация текста выполняется с помощью экспериментально установленных типовых, структурных схем, которые задаются операторами идентификации и ключевыми словами в проверочных моделях. Под "словом" в проверочной системе ЭЛИС и в дальнейшем тексте этой статьи понимается любое число (в пределах поля ответа) печатаемых символов, в том числе и любой один знак (например, знак препинания). То есть это некоторый элемент текста, не всегда соответствующий лингвистическому пониманию слова. Степень детализации определяется экспертными знаниями и задачами составителя заданий и позволяет регулировать допустимую, по его мнению, точность и неточность возможных ответов в пределах сохранения правильного смысла ответа. Набор операторов образует логическую часть проверки, ключевые слова – фактическую. Такое сочетание общеизвестно, но способы и структура представления проверочных моделей в экспертной системе ЭЛИС, которые можно использовать обычным пользователям компьютера, позволили получить одно из возможных решений задач автоматизированной идентификации текста для ППС.

Таким образом, автоматизированная проверка текстовых ответов в инструментальной системе тестирования ЭЛИС, как и во многих других подобных системах, выполняется сравнением логических моделей позиций ключевых слов, заданных составителем тестовых заданий, с выделяемой из текстового ответа логической моделью позиций слов, найденных в ответе испытуемого. В тоже время, система ЭЛИС имеет свои оригинальные решения, позволяющие программировать автоматизированную проверку более сложных текстовых ответов по сравнению с возможностями известных версий инструментальных систем тестирования (SuperTest, SunRav TestOfficePro и др.). В экспертной системе ЭЛИС не используются общепринятые подстановочные знаки типа "*", "?", эффективные для простого поиска текста, но затрудняющие читаемость сложных проверочных моделей, приводящие к неестественному виду текста для составителей тестовых заданий. В системе идентификации текста ЭЛИС экспертные знания накапливаются в хорошо обозримых таблицах, образующих иерархию проверочных моделей. Эти таблицы можно заполнять в текстовом виде или с помощью мастера проверочных моделей.

Экспертная система принятия решений о правильности ответа испытуемого на основе проверочных моделей состоит из трех уровней, управляется операторами и параметрами логических моделей ответов составителей тестовых заданий.

вых заданий (Рис.1). Каждый уровень имеет произвольное, задаваемое составителем, множество входных точек (результаты анализа в подчиненных системах) и один выход (решение). Верхний уровень образует текущая совокупность проверочных моделей связываемых отношениями AND, OR, NOT. По своему строению такие проверочные модели каждого задания текущего теста неотличимы от моделей из общего банка моделей всей системы тестирования ЭЛИС. Объем каждой модели, общее число используемых моделей определяется потенциальной сложностью проверки ответа текущего задания. Второй уровень образует каждая текущая модель из совокупности проверочных моделей одного задания. В ее состав входят операторы третьего уровня проверки, относящие каждое проверочное слово к одной из схем групп слов и схем порядков слов или выполняющие обращение к внешнему, пополняемому банку общих моделей. Отдельные таблицы общего банка проверочных моделей вызываются из любой текущей модели любого задания.

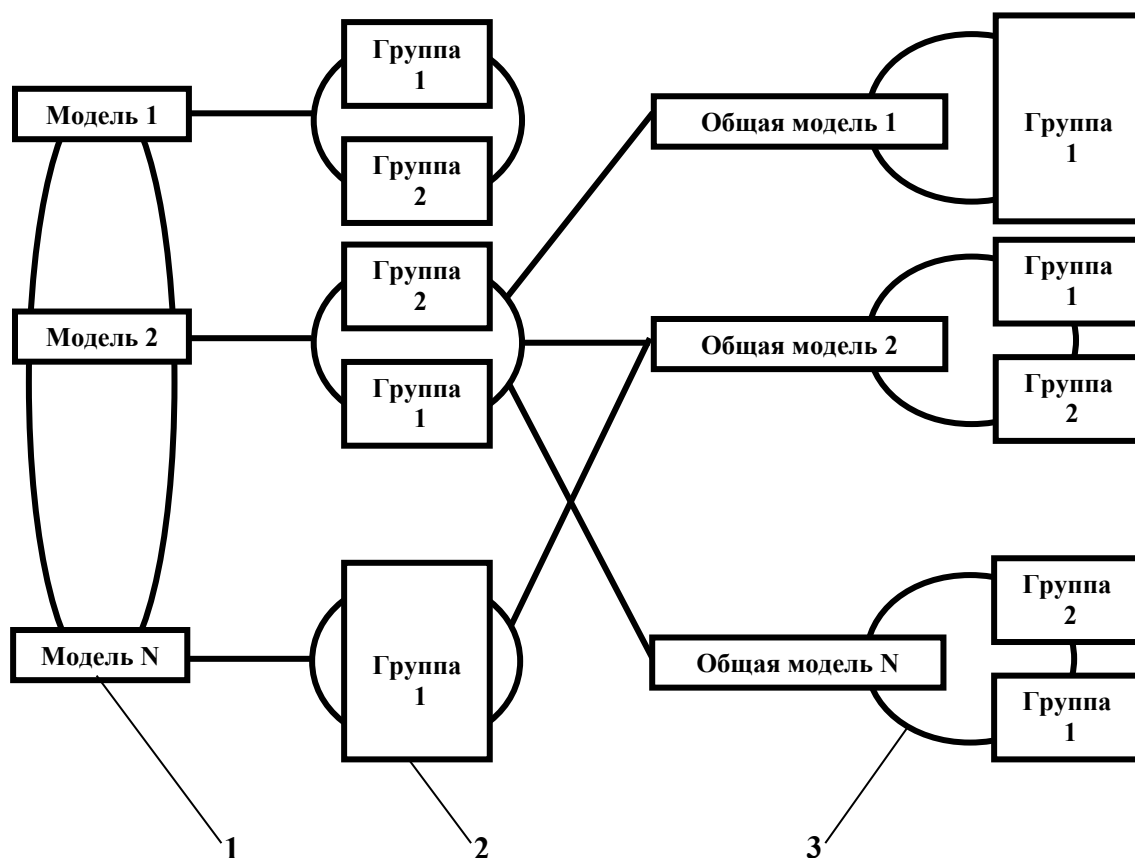


Рис.1. Общая схема моделей идентификации для одного ответа

- 1 – Внешний, основной слой набора моделей (1 и 2 уровень).
- 2 – Группы слов в моделях проверки возможных ответов (3 уровень).
- 3 – Общий банк дополнительных моделей.

Эти же операторы (как говорят – два в одном) относят каждое из слов к основному слову, к элементу текста эквивалентному основному слову (часто это синоним в лингвистическом понимании текста), к недопустимому элементу текста (часто это антоним в лингвистическом понимании текста). Другие – задают интервалы, виртуальность слов, рефлексы. В определении интервалов поиска следующих слов относительно найденного текущего слова в проверочной

системе ЭЛИС используются позиция и длина слова, дистанции в знакоместах, слитность слов или произвольность дистанции. Виртуальными называются слова, для которых, как и для других слов ответа проверяется логика позиций, но их отсутствие в тексте ответа не приводит к отрицанию идентификации. Виртуальные элементы, создавая избыточность проверки, позволяют использовать для анализа ответов нечеткие проверочные границы.

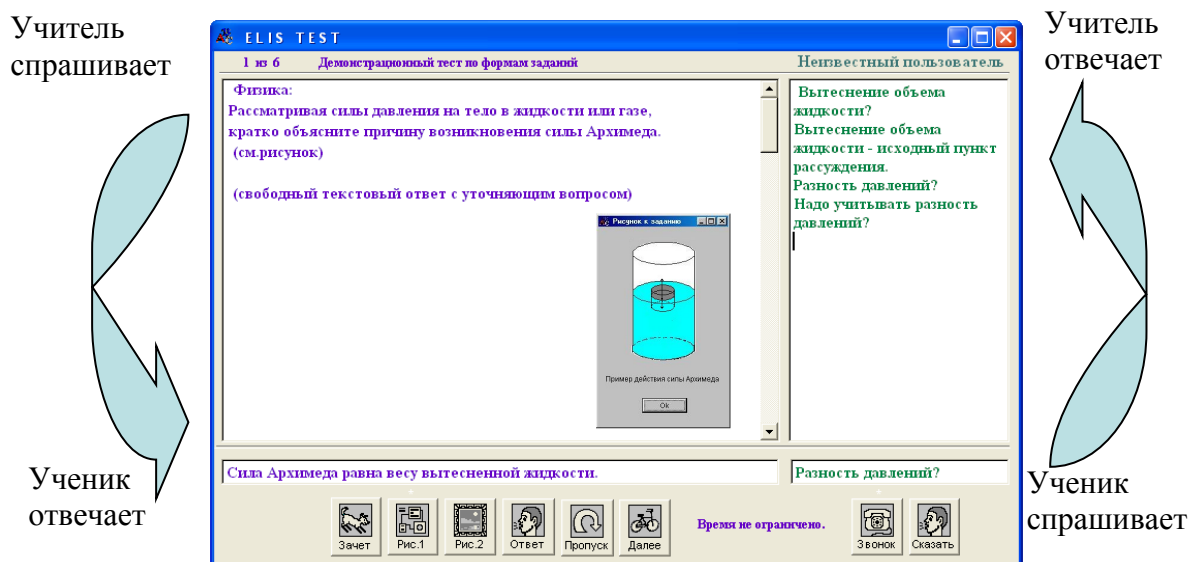


Рис.2. Проверочный интерактивный диалог в режиме самообучения

В проверочных моделях используется позиционная логика самих операторов, что позволяет наиболее компактно и наглядно задавать параметры проверки для идентификации смыслового множества допустимых позиций слов в текстовых ответах. Сходимость к смысловой правильности ответа обеспечивается несколькими путями. Во-первых, это специальные условия и ограничения в тексте задания (вопроса) теста. Во-вторых, программируемые рефлексы (реакции проверочной системы) на слова и рефлексы на модели позиций слов, обнаруживаемые в тексте ответа испытуемого и объявляемые во время проверки. В-третьих, это проверка в форме диалога с испытуемым (Рис.2) и возможным выходом на уточняющий (в пределах проверочного континуума) однозначный ответ по типу выборочного, кроссвордного или координатного ответа на рисунке. Диалоговая форма используется так же и для усовершенствования экспертных проверочных моделей с помощью создаваемого в процессе проверки банка ответов и протоколов проверки. С этой же целью используются статистические методы анализа результатов проверки: распределение заданий по группам – факторам проверки, автонастройка весовых коэффициентов трудности заданий и автоопределение коэффициентов сложности заданий.

Группы, порядки, интервалы поиска, рефлексы, виртуальные слова, эквивалентность слова и модели позволяют задавать составителю тестовых заданий смысловую точность идентификации, устанавливать правильность ответов и/или наличие знаний у испытуемого. Это основные средства программирования проверки ответов в используемом анализаторе текстов. Экспериментально

изучались и другие решения по оценке взаимного расположения позиций слов и интерпретации их значений. В анализаторе текстов ЭЛИС найдены способы построения произвольной внутренней системы проверочных моделей для организации параллельной обработки текста, получено наиболее компактное и близкое к естественному виду программирование сравнения текстов. В данном исследовании установлено, что наибольшей внешней наглядностью и простотой восприятия для сложных проверок обладают не громоздкие выражения с использованием операторных скобок, а компактные модели на основе иерархии таблиц, структурой которых задается синтаксис языка проверки ответов.

К преимуществам экспертной системы относится возможность развития компьютерных проверочных моделей до состояния эффективной проверки в типовых проверочных ситуациях. К недостаткам: невозможность анализа заранее не прогнозируемых сообщений. В то же время накопление экспертных знаний позволяет повысить уровень эвристичности анализа текстовых ответов. Эффективность предлагаемых решений для использования в разработке ППС основывается на том, что экспертная система проверки текстовых (на этой основе возможна проверка выборочных и других ответов) в инструментальной системе ЭЛИС согласуется с механизмами стандартизации тестовых баллов, факторной диагностикой уровня знаний испытуемых, механизмами экспорта и импорта заданий, средствами защиты тестовой системы и другими подсистемами. Рассмотрение некоторых педагогических особенностей и технологических приемов идентификации текстов на примере инструментальной тестовой системы не исчерпывает этого направления в исследовании ППС. Потребность в экспертных методах для учебного процесса весьма высока. Это не только автоматизация контроля знаний, но и поисковые задачи для электронных учебников и энциклопедий, интеллектуализация электронных репетиторов, составление индивидуальных графиков обучения и многие другие задачи.

Публикации автора по теме исследования:

Сравнение текстов по моделям позиций слов/ журнал "Информатика и образование", 1990, №4. ISSN 0234–0453. –С.45-49.

Логические модели проверки ответов в компьютерных педагогических тестах// XII конференция-выставка "Информационные технологии в образовании". Сборник трудов участников конференции. /ИТ в контроле и оценке результатов обучения. Часть V. –М.: МИФИ, 2002. –С.25-27

Experimental designing of the language of text identification within the Instrumental Testing System ELIS. /Information and Communication Technologies and the Knowledge Economy. Volume 1 //eAdoption and the Knowledge Economy Issues, Applications and Case Studies. Edited by Paul Cunningham and Miriam Cunningham. Part 2. –Amsterdam, Berlin, Oxford, Tokyo, Washington, DC: IOS Press, 2004. pp 1101-1107.