

# **МЕТОДОЛОГИЯ ФОРМАЛИЗАЦИИ ЗНАНИЙ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ ОБРАЗОВАТЕЛЬНОГО НАЗНАЧЕНИЯ**

**Шаров А.Г.**

Современные потребности экономики и бизнеса в трудовых ресурсах требуют сертификации специалистов как неотъемлемой процедуры, подтверждающей качество их профессиональной подготовки и профессиональной пригодности. При общении друг с другом профессионалы используют специфические термины своей предметной области, не давая им расшифровку. Знание терминологии своей предметной области и понимание её смыслового содержания является важной составляющей профессиональной подготовки специалистов. Ещё на этапе обучения эти специалисты должны научиться правильно воспринимать и воспроизводить смысл этих терминов. Любой процесс обучения завершается проверкой знаний обучаемого, включающей контроль знания обучаемым терминологии изучаемой дисциплины, предметной области.

В то же время, изменяющиеся концепции естествознания, отдельных предметных областей неизбежно приводят к семантическим изменениям терминологической базы – тезауруса предметной области. Отсюда следует вывод о необходимости корректировки учебного материала, дидактического, лекционного, и, соответственно, баз знаний автоматизированных обучающих систем.

Вместе с тем, корректировка методических и дидактических материалов в соответствии с современными требованиями к качеству профессиональной подготовки, зачастую ведется методом проб и ошибок, а используемый российскими учебными заведениями набор приёмов методологического воздействия на качество обучения достаточно ограничен. Учебные заведения испытывают острую потребность в эффективном автоматизированном инструментарии методологического сопровождения учебного процесса для проведения адекватных изменений формы и структуры учебного материала, направленных на его совершенствование.

Существующие автоматизированные обучающие системы имеют принципиальное ограничение: системы осуществляют контроль корректности текстового ответа, используя механизм полного поэлементного совпадения введенной текстовой фразы с эталонным текстом. В результате контроля система формирует бинарный выходной сигнал о совпадении текстов, что является реализацией способа управления обучением, построенного на бинарных деревьях решений. С позиций автоматизации методологического воздействия на корректировку методик обучения данный механизм является достаточно примитивным, так как не обеспечивает необходимой полноты анализа знаний обучаемых, и, соответственно, может ограниченно применяться

лишь для управления последовательностью предъявления обучаемому информационных материалов либо тестовых заданий посредством визуального интерфейса системы. Кроме того, ни одна из известных обучающих систем не осуществляет автоматический анализ текста произвольного содержания, введенного обучаемым в процессе тестирования. Данный недостаток значительно снижает возможности использования этих систем в плане контроля знаний обучаемых, например, усвоения терминологии изучаемой предметной области.

Система контроля корректности текста СККТ имеет, по сравнению с другими системами, повышенные методологические и дидактические характеристики, поскольку позволяет произвести не только качественную, но и количественную оценку произвольного текста, вводимого обучаемым при проведении тестирования. Это предоставляет возможность не только проведения более детального анализа и формирования методологически корректного воздействия на последовательность предъявления учебного материала, но и корректировки содержания учебного курса (дидактического материала, тезауруса и словарей предметной области). Универсальность данной модели позволяет адаптировать её практически к любой предметной области, имеющей свой профессиональный тезаурус, являющийся терминологической базой профильного обучения. Адаптация производится простым подключением словарей другой предметной области.

Предлагаемый подход к решению задачи интеллектуализации автоматизированных систем контроля знаний основан на предположении о релевантности двух текстовых фраз произвольной длины и подразумевает вычисление релевантности (количественных характеристик соответствия ответа эталону) на основе количественных показателей синонимии терминов предметной области. Методы обработки ответа и вычисления релевантности могут быть различными и зависеть от интерпретации задачи.

Например, необходимо сопоставить каждой формулировке термина предметной области (ответу обучаемого) ее образ-вектор (ответ эксперта, эталонное определение) в некотором пространстве таким образом, чтобы вектора сходных формулировок были близки. Эта задача относится к задачам классификации. Для ее решения целесообразно использование алгоритмов функционирования нейроподобных структур. Это связано с тем, что такие алгоритмы лучше остальных зарекомендовали себя при решении задач классификации и кластеризации данных. В частности, нейроподобная структура, основанная на последовательном применении сети динамической ассоциативной памяти и модифицированной сети Хемминга, как показывают модельные испытания, помогает решить поставленную задачу.

Основой другого решения поставленной задачи является модификация векторной модели текста, предложенной Дж. Солтоном. Способность классифицировать бинарные вектора, которую предоставляют нейронные сети, хорошо сочетается с классической моделью Солтона.

Не вдаваясь в дискуссию о том, что в настоящее время можно считать базой знаний в автоматизированной обучающей системе, для наших целей будем называть базой знаний в общем случае словарь предметной области, содержащий упорядоченный определённым образом набор терминов, связанных между собой отношениями синонимии, обладающими количественными характеристиками – так называемыми показателями синонимии. Для целей практической реализации автоматизированной системы контроля знаний на основе предложенных подходов, придерживаясь основополагающего принципа – количественных показателей синонимии, в состав базы знаний системы должны входить следующие компоненты:

- тезаурус предметной области – толковый словарь терминов и их эталонных определений;
- словарь синонимов терминов предметной области;
- словарь служебных слов, специальных символов и стоп-слов (общеупотребительных слов, не влияющих на семантику текста).

Суть предлагаемого автором метода заключается в экспертном заполнении словарей и таблиц синонимов начальными значениями коэффициентов синонимии и последующей их адаптацией к устойчивому состоянию посредством нейроподобной структуры, основанной на последовательном применении сетей различной топологии, что, как показывают модельные испытания, помогает решить поставленную задачу. На основании экспериментальных данных результатов работы системы контроля знаний, реализованной на выше изложенных принципах, можно сделать вывод о возможности автоматической корректировки баз знаний (тезауруса и таблиц синонимии), используемых при изучении определённой предметной области.

Основным методологическим приёмом верификации (проверки достоверности) корректировки базы знаний является превентивное изменение эталонной либо экспертной формулировки терминов предметной области. Повторное тестирование с использованием системы контроля корректности текста при этом даёт новые числовые показатели, стабилизирование (выравнивание) значений которых в отдельности по группе экспертов–преподавателей и обучаемых свидетельствует, с одной стороны, о корректно вычисленных нейроподобной структурой коэффициентах синонимии, и, с другой стороны, о более приемлемой для понимания студентами формулировке термина в тезаурусе предметной области.

Одним из отличий от ранее разработанных систем является возможность использования существующих баз данных и баз знаний как непосредственно содержательной части курса обучения независимо от расположения, места нахождения знаний (архивов, баз, банков), а также конкретного формата их представления в виде электронного документа, для поиска синонимов и формирования количественных показателей синонимии терминов.