

ТЕСТИРОВАНИЕ НА ОСНОВЕ СЕМАНТИЧЕСКОЙ БАЗЫ ЗНАНИЙ

Системы тестирования обычно используют режим выбора тестируемым правильного варианта из предложенного списка вариантов для ответа слов, ситуаций, схем и т. п. Если же вопросы теста предполагают ответы на естественном языке в виде развернутого предложения, то их обработка, как правило, производится «вручную» экспертом-преподавателем, что существенно усложняет и замедляет процесс подсчета результатов тестирования.

Современные технологии позволяют разрабатывать компьютерные обучающие системы, автоматизирующие процесс обучения с заменой реального преподавателя виртуальным и функционирующие в режиме диалога с обучаемым по учебному материалу изучаемой дисциплины на естественном языке [1]. Для организации такого диалога для текстовой информации по данной предметной области создается база знаний с семантической структурой, которая содержит все ключевые понятия-объекты и их смысловые связи по учебному материалу дисциплины и функционально представленная в виде семантической сети [2].

Структура семантической базы знаний может быть представлена в виде записей-триад и храниться в двумерном текстовом массиве из отношений между ключевыми понятиями для каждого простого предложения в тексте учебного материала в виде семантической связи: «субъект»-«отношение»-«объект» или в лингвистических терминах – в виде структурных блоков предложения: «подлежащее»-«сказуемое»-«дополнительные члены предложения», которая позволяет использовать заложенные в триадах семантические связи для интеллектуальной обработки текстовой информации.

Такую структуру семантической базы знаний можно использовать как для изучения дисциплины в режиме диалога с задаваемыми вопросами по учебному материалу и получаемыми на них точечными ответами из базы знаний [3], так и для проведения тестов на естественном языке. Например, для фразы «Электрон имеет отрицательный заряд» на тему учебного материала по электричеству в двумерном текстовом массиве базы знаний будет храниться запись-триада:

«Электрон»-«имеет»-«отрицательный заряд»
--

Для проведения теста на естественном языке добавим для удобства в эту строку двумерного массива четвертым элементом соответствующий ей вопрос:

«Какой заряд имеет электрон?»

который будет выводиться на экран в процессе тестирования. Пусть в качестве ответа на этот вопрос тестируемый введет фразу, например:

«Электрон имеет положительный заряд»

что будет являться неправильным ответом.

Для проверки полученного ответа это предложение разбивается на отдельные слова и раскладывается на триаду, передающую семантическое отношения между субъектом и объектом в этом предложении:

«Электрон»-«имеет»-«положительный заряд»

Чтобы образовать такую триаду из отдельных слов в ответном предложении необходимо, прежде всего, найти среди всех слов предложения сказуемое, которое определяется по псевдоокончанию, характерному для слов кандидатов на сказуемое. В приведенном примере ответного предложения это будет псевдокончание «ет» в слове «имеет». Тогда слева от сказуемого, как элемента триады «отношение», будет элемент триады «субъект», а справа – «объект».

Затем необходимо пройти в цикле все записи в массиве базы знаний и найти среди них строку, содержащую соответствующую запись-триаду с верной информацией.

При определении соответствия элементов триады ответа и триады в базе знаний в общем случае необходимо использовать шаблонные регулярные выражения для всех элементов триады полученного в ответе предложения, поскольку в правильном ответе подлежащие могут стоять в других падежах, а сказуемые в других склонениях, чем в верной записи базы знаний, но их необходимо засчитать как правильные. Однако при проведении тестирования достаточно будет проверить на соответствие только два элемента триады: «отношение» и «объект», поскольку первый элемент – «субъект» уже содержится в задаваемом вопросе и должен будет по необходимости использоваться в ответном предложении по принятым в системе тестирования правилам его построения.

В языке JavaScript для обработки регулярных выражений обычно используются экземпляры из класса `RegExp()`, аргументом в которых выступает заданный шаблон исходного выражения. При этом, если блок подлежащего состоит из нескольких слов, то их необходимо объединить в одну последовательность символов, используя связку символов «.*», для каждого слова. Например, для «объекта» в триаде

приведенного выше ответного предложения шаблоном будет последовательность символов «.*положительный.*заряд.*», который после проверки в системе тестирования даст отрицательный результат. Таким образом, шаблон для подлежащих в ответном предложении определяется как последовательность символов, задаваемых на проверку их соответствия в верном выражении для блока подлежащего в базе знаний. Причем проверка шаблона будет выполняться для количества символов в нем больше трех, поскольку символы в шаблоне «.*» не нуждаются в проверке.

Шаблон для сказуемого в ответном предложении строится на основании таблицы псевдоокончаний сказуемых, которая используется для замены конкретного окончания сказуемого в ответе на набор всех возможных окончаний для этого сказуемого в различных спряжениях первого и второго рода. Например, для построения шаблона сказуемого «имеет» необходимо в функцию `RegExp()` вложить шаблон «име|ет|ут|ют» и ответ типа: «Электроны имеют отрицательный заряд» будет зачитан как правильный.

Для проверки соответствия введенного тестируемым ответа на заданный вопрос система тестирования проходит в цикле все записи в базе знаний и проверяет на выполнение сложного условия из связки элементов триады «отношение» и «объект» в ответе каждой записи в базе знаний, используя при этом метод `test()` для обработки регулярных отношений. Причем результат проверки системой каждого вопроса выводится на экран фразой «ПРАВИЛЬНО!» при получении правильного ответа, либо фразой «НЕПРАВИЛЬНО!» при получении неправильного ответа. После последовательного прохождения всех вопросов теста общий результат выдается на экран и подсчитывается по формуле:

$$\boxed{\text{«ВАШ ОБЩИЙ РЕЗУЛЬТАТ»} \Rightarrow m \times 10 / n \text{ «баллов»}}$$

где m – количество правильных ответов, а n – количество заданных в тесте вопросов.

Таким образом, семантические отношения между учебными объектами в базе знаний позволяют провести анализ результатов тестирования не только по точному совпадению ответа с соответствующей записью в базе знаний, но и по совпадению их смыслового содержания, заложенного в шаблонах для «отношения» и «объекта», используемых при проверке соответствия ответа и верной информации, заложенной в триадах базы знаний. При этом система тестирования обеспечивает вывод вопросов для тестирования на естественном языке в виде текста либо голосом с последующим анализом ответа тестируемого, введенного им также в виде текста либо голоса, используя в

последнем случае для диалога подключение к речевым сервисам для распознавания речи в текст и синтеза речи из текста.

ЛИТЕРАТУРА

1. Гурин, Н.И. Компьютерные обучающие системы в издательском деле / Н.И. Гурин, О.В. Герман. – Минск: БГТУ, 2015. – Ч. 2 – 192 с.
2. Гурин, Н.И. Генератор семантической сети информационной системы в таблицу реляционной базы данных / Н.И. Гурин, Я.А. Жук // Труды БГТУ. – 2015. – № 6: Физ.-мат. науки и информатика. – С. 181–185.
3. Жук, Я.А. Реализация диалога с компьютерной обучающей системой на языке JavaScript с помощью веб-сервисов / Я.А. Жук, Н.И. Гурин // Труды БГТУ. – 2018. – № 6: Физ.-мат. науки и информатика. – С. 190–196.

УДК 658.512.6

В.А. Алешаускас, ассист.
(БГТУ, г. Минск)

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ АВТОМАТИЗАЦИИ УПРАВЛЕНИЯ ИЗДАТЕЛЬСКО-ПОЛИГРАФИЧЕСКИМ КОМПЛЕКСОМ С ПОМОЩЬЮ КАНБАН

Канбан – метод управления производством, реализующий принцип «точно в срок» и способствующий равномерному распределению нагрузки между работниками. При данном подходе весь процесс прозрачен для всех членов команды. В системе производства показано, что необходимо создать, когда и сколько. Задачи по мере поступления заносятся в отдельный список, откуда каждый исполнитель может извлечь требуемую задачу. Метод основан на одноименном методе в производственной системе «Тойоты» [1].

Метод «канбан» принято использовать как инструмент распределения задач при разработке программного обеспечения с использованием виртуальной либо реальной доски и стикеров с задачами на ней. Данный метод производства был внедрен в частном издательско-полиграфическом комплексе, который специализируется на издании книг малыми тиражами, ООО «Колорград» (издательский дом «Сегмент»). В начале внедрения метода управления на предприятии уже использовалась CRM-система, которая организовывала работу менеджеров по работе с клиентами предприятия и их заказами. В момент, когда появилась необходимость автоматизации производства, выбор