

УДК 811.161.1'243:378.147.168

ИНТЕЛЛЕКТУАЛЬНЫЙ ЛИНГВОПРОЦЕССОРНЫЙ КОМПЛЕКС «КЛИОС» ДЛЯ ОБУЧЕНИЯ РКИ

Горисев С.А.¹, Койнов А.В.¹, Куземчик В.Д.¹, Лисинин С.В.¹, Михалева Е.В.¹, Мишунин О.Б.¹, Савинов А.П.¹, Терехин Д.Э.¹, Фирстов Д.И.¹, Черкашин А.Ю.¹

¹ФГБОУ ВПО Национальный исследовательский Томский политехнический университет, Томск, Россия (634050, пр. Ленина, 30), e-mail: lisininsv@gmail.com

В данной статье рассмотрены автоматизированные обучающие системы, применяемые сегодня в России для создания электронных обучающих курсов по различным дисциплинам. Проанализированы характерные особенности данных систем и выявлены их недостатки на примере построения курсов для изучения русского языка как иностранного. Проведен анализ требований для разработки интеллектуальной обучающей системы нового поколения, в основе которой лежит лингвистический процессор, позволяющий отойти от классических методов тестирования и оценивать развернутые ответы обучаемого на естественном языке. На основе потребностей конечных пользователей и, исходя из современных достижений в области искусственного интеллекта, был частично реализован тренажер «КЛИОС» с модулем автоматического решения упражнений раздела русской грамматики «Система именного склонения». Тренажер оснащен подсистемами, призванными сделать процесс обучения языку более комфортным и эффективным. «Подсказки» помогают лучше понять суть упражнения и корректно выполнить его. Автоматическое формирование ответов на естественном языке подсистемой «Проверки упражнений» позволяет обучаемым быстро проанализировать допущенные ими ошибки. Преподаватель имеет удобный инструмент для разработки упражнений и автоматически заполняющийся электронный «Журнал успеваемости» его групп. Рассмотренная система проходит апробацию в учебном процессе.

Ключевые слова: интеллектуальная обучающая система, лингвистический процессор, морфологический анализатор, синтаксический анализатор.

INTELLIGENT TUTORING SYSTEM “KLIOS” BASED ON THE LINGUISTIC PROCESSOR FOR STUDING RUSSIAN AS A FOREIGN LANGUAGE

Gorisev S.A.¹, Koynov A.V.¹, Kuzemchik V.D.¹, Lisinin S.V.¹, Mikhaleva E.V.¹, Mishunin O.B.¹, Savinov A.P.¹, Terekhin D.E.¹, Firstov D.I.¹, Cherkashin A.Y.¹

¹National Research Tomsk Polytechnic University, Tomsk, Russia (634050 Tomsk, Lenin prospect, 30), e-mail: lisininsv@gmail.com

In this article describes the automated tutoring systems for e-learning courses development which are used in Russia today. The characteristics of these systems were analyzed and, based on the example of Russian as foreign language syllabus development, shortcomings were identified. The requirements for the intelligent tutoring system of a new generation development were analyzed. The linguistic processor is laid in its foundations and allows to get rid of the classical methods of testing; also it helps to evaluate student's answers in natural language. The developing training system is based on the needs of end-user groups, achievements in artificial intelligence and telecommunications network development. It is partially implemented as a “KLIOS” system shell with integrated algorithms of automatic exercise Russian language section “Nominal declension”. The tutoring system is equipped with subsystems designed to make learning the language easier and more efficient. “Tips” help students understand the exercises better and correctly execute them. Automatic generation of natural language answers by the subsystem “Exercise checker” allows to analyze mistakes quickly. The teacher has a handy tool for developing exercises and automatically filling electronic “Gradebook” of his groups. “KLIOS” is currently being piloted in the educational process.

Keywords: intelligent training system, linguistic processor, morphological analyzer, syntactical analyzer (parser).

В Томском политехническом университете обучается русскому языку значительное число иностранных граждан (в прошлом году более 900 человек), прибывших в университет для

освоения различных образовательных программ. Изучение русского языка иностранными гражданами связано с рядом трудностей. В первую очередь это касается собственно языковых трудностей: сложность самого русского литературного языка (богатая грамматическая система, большое количество флексий, развитая лексическая система, различные виды омонимии и полисемии и т.п.), необходимость осваивать профессиональный дискурс. Кроме того, более 30% обучающихся в ТПУ иностранных студентов – представители Азиатско-Тихоокеанского региона и, следовательно, сложности освоения принципиально иной языковой системы проявляются на всех ее уровнях: фонетика и графика, морфемика и словообразование, лексика, морфология и синтаксис. Помимо лингвистических, возникают определенные трудности психологического характера: сложности языковой, академической и социокультурной адаптации в инокультурной среде, различный уровень знаний на момент начала обучения, разные когнитивные стили и черты характера, а также особенности восприятия и переработки информации, характеризующие нынешнее молодое поколение: преимущественно визуальный канал получения информации, ее гипертекстуальное оформление. Есть сложности и в организации процесса обучения: различные сроки заезда, формирование групп с различным уровнем языковой адаптации. Все это приводит к тому, что возникает острая необходимость в создании обучающих ресурсов, которые были бы ориентированы на индивидуальные стратегии обучения, на привычные формы получения информации, на возможности дистанционного обучения, в общем – помогли бы современному студенту в его адаптации и выборе оригинальных стратегий обучения.

Освоение русского языка «с нуля» требует еще и дополнительных условий для выработки автоматизма в использовании языковых средств того или иного уровня.

В этих условиях особое значение приобретает создание компьютерной системы, обеспечивающей возможность организации самостоятельной работы студентов, отвечающей принципам индивидуализации обучения и ресурсоэффективности (экономии временных, материальных и человеческих ресурсов). Такая система должна не только дать определенную «свободу» обучающимся, но и «разгрузить» преподавателя, взяв на себя ряд функций по проверке и корректировке стратегии обучения.

В ходе анализа процесса обучения иностранцев в Томском политехническом университете были сформированы функциональные требования к системе, выполнение которых должно обеспечить эффективность ее использования:

1. Система должна обладать справочно-обучающими функциями и предоставлять студентам удобный доступ к учебным материалам по любой тематике.
2. Система должна выполнять функции тренажера, позволяя студенту самостоятельно отрабатывать полученные языковые навыки посредством выполнения упражнений.
3. Система должна выполнять контролирующую функцию, оценивая знания студента по множеству критериев: орфография, словообразование, синтаксис, пунктуация, понимание смысла текста и пр.
4. Система должна автоматически обеспечивать индивидуализацию, т. е. используя актуальную информацию о знаниях студента, адаптировать индивидуальный план таким образом, чтобы основной упор делался на проблемных местах.
5. Как следствие предыдущего пункта, система должна «помогать» студенту по ходу выполнения тренировочных заданий, предоставляя грамматические правила по разделам, вызывающим затруднения, перевод на родной язык, проверку орфографии и т. п.
6. Система должна быть доступна для использования через Интернет.

На сегодняшний день на российском рынке обучающих систем наибольшее распространение получили три: «Blackboard» [7], «Hot Potatoes» [8] и «Moodle» [4]. В целом возможности для разработки учебных курсов в рамках данных систем незначительно отличаются друг от друга: все материалы и тесты размещаются в них на двух независимых подсистемах: обучающей и контролирующей.

Подсистема обучения представляет собой программу, осуществляющую вывод в установленной последовательности порций учебного материала в гипертекстуальной форме с использованием мультимедийных средств.

Контролирующая подсистема применяется для проведения тестирования с использованием методов стандартизованного контроля знаний. Их сущность состоит в том, что обучаемому предлагается выборка заданий и по результатам выполнения тестов выносится суждение об уровне успешности прохождения курса. Ответ классифицируется как правильный, неправильный или частично правильный (например, неполный) путем его сравнения с заранее указанным эталоном. С помощью такого инструмента очень сложно автоматизировать оценку знаний по различным критериям. Конечно, можно составить тестовые упражнения на каждый из приведенных в п. 3 критериев, но это приведет к резкому увеличению трудоемкости создания обучающих курсов. Гораздо более эффективным и естественным подходом является комплексная оценка каждого выполненного студентом задания.

Одним из самых больших минусов [7], [8], [4] является отсутствие функционала, который позволял бы автоматизировать адаптацию учебного плана, в зависимости от качества усвоения материала обучаемым. Таким образом, рассмотренные выше инструментальные комплексы не обладают средствами, необходимыми для разработки полноценных курсов по изучению русского языка.

Однако существует и иное направление в компьютерном обучении – интеллектуальные обучающие системы (ИОС) [1]. В России работы по созданию ИОС ведутся на данный момент в сравнительно небольших масштабах и пока еще не вышли на коммерческий уровень. К числу пионерских работ в области создания интеллектуальных обучающих систем можно отнести проект ИОС «Волга». Она позволяет осуществлять углубленную индивидуализацию обучения на основе логико-оптимизационных методов автоматизации планирования учебного процесса, а также использует оригинальные методы дедукции, абдукции, решения логических уравнений и многокритериальной оптимизации [5]. Эта система направлена на обучение предметам естественнонаучного цикла. Полноценных отечественных аналогов, разработанных для других дисциплин, не обнаружено.

Два этих фактора – потребность в высокоэффективных обучающих системах и их неразработанность – привели к тому, что в ТПУ было принято решение самостоятельно создать интеллектуальный лингвопроцессорный тренажер обучения РКИ, который удовлетворял бы сформулированным выше требованиям.

Чтобы описать принцип работы тренажера, приведем примерную структуру типового упражнения. Большинство из них состоят из задания и списка однотипных вопросов. В задании часто приводится пример вопроса и правильного ответа на него (так называемая «модель упражнения»). Чаще всего в заданиях от студента требуется дать развернутый ответ на поставленный вопрос. Например:

Вопрос: *Где вы были вчера? (театр)*

Ответ: *Я был вчера в театре.*

Подобные упражнения позволяют тренировать языковые навыки комплексно: для правильного ответа на вопрос требуется и знание словообразования, и знание пунктуации, и понимание смысла текста или фразы. В то же время проверку подобных упражнений очень нелегко автоматизировать по двум причинам. Во-первых, в русском языке относительно свободный порядок слов: можно сказать «**Вчера** я был в театре», «Я был в театре **вчера**» и т. д. Во-вторых, ответ может меняться в зависимости от пола отвечающего («был/была») и

смысла, который он вкладывает в вопрос (можно предположить, что вопрос задан группе и ответить «*Мы были вчера в театре*»). Таким образом, при использовании обычных для компьютерного обучения методов контроля знаний, преподавателю придется при создании упражнения на каждый вопрос готовить значительное число правильных ответов. Например, для данного вопроса пришлось бы приготовить минимум девять возможных ответов, а для упражнения, состоящего из десяти подобных вопросов, – около 90 ответов.

Чтобы снизить трудоемкость подготовки упражнений и сделать возможной заявленную в требованиях многокритериальную оценку студенческих ответов, в системе «КЛИОС» используется лингвистическое программное обеспечение: библиотеки морфологического анализа и склонения слов *Руморфу 2* [9], а также синтактико-семантический анализатор *АВВУУ Comprero* [6]. Благодаря этим средствам, становится возможным следующий сценарий проверки упражнения:

1. Синтактико-семантический анализ примеров вопроса и ответа. Построение на его основе алгоритма генерации ответов.
2. Применение полученного алгоритма к вопросам упражнения. Построение синтаксического дерева правильного ответа (учитывающего все возможные варианты) для каждого вопроса упражнения.
3. Синтактико-семантический анализ ответов студента и построение их синтаксических деревьев; сравнение данных деревьев с выходными данными пункта 2:
 - сравнение отдельных слов, выявление орфографических ошибок и ошибок словообразования;
 - сравнение деревьев целиком, выявление синтаксических ошибок и ошибок в выполнении задания;
 - сравнение текстового представления ответов, определение пунктуационных ошибок и ошибок, связанных с неверным порядком слов.

На основе выявленных ошибок происходит заявленная в требованиях корректировка курса. В базе данных собирается статистика ошибок каждого студента. Если обучаемый часто делает ошибки в определенных словах, система будет выбирать упражнения, в которых встречаются проблемные слова. Если затруднения вызывают конкретные синтаксические структуры, то упражнения с этими структурами будут также чаще попадаться. Если у студента выявлена проблема с освоением какой-либо темы, то система просто увеличит количество упражнений по ней.

Для того чтобы процесс обучения шел эффективней, студенту, испытывающему трудности с тем или иным заданием, система «КЛИОС» может оказывать некоторую помощь. Во-первых, она может указать на орфографические ошибки до проверки упражнения и дать возможность их исправить.

Во-вторых, в «КЛИОС» присутствует грамматический справочник, представляющий собой системно организованный в гипертекстуальном формате справочный ресурс. При этом обучаемый имеет быстрый доступ к нужному материалу непосредственно в процессе выполнения задания.

В-третьих, в упражнениях имеется быстрый доступ к ресурсу Google Translate – бесплатной службе, позволяющей пользователю мгновенно перевести текст на понятный ему язык без необходимости использования каких-либо дополнительных словарей в самом тренажере. Интерфейс системы предоставляет студенту возможность выделить непонятное слово или фразу и сразу же получить перевод на один из 64 языков [3] прямо в процессе выполнения упражнения.

Перечисленные вспомогательные функции призваны сделать обучение максимально комфортным за счет объединения в одном месте всех необходимых ресурсов, которые могут понадобиться в процессе освоения РКИ. Их количество будет увеличиваться за счет внедрения новых программных модулей.

Таким образом, у студентов появляется заявленная в требованиях возможность заниматься с системой самостоятельно, снижается порог вхождения в образовательную программу, уменьшается загрузка преподавателей, задача которых – только создавать упражнения и уроки по отдельным темам русского языка, а также следить за ходом обучения студентов. Для этого в их распоряжении есть следующие инструменты:

1. Редактор упражнений. Позволяет создавать тренировочные задания за несколько простых шагов. При этом от преподавателя не требуется знаний языков программирования и разметки.
2. Редактор уроков. Этот инструмент позволяет преподавателю создавать списки упражнений по определенным темам русского языка. При этом каждый урок может быть назначен как целой группе, так и конкретным студентам. Срок прохождения занятия можно ограничить.
3. Контроль успеваемости обеспечивается двумя инструментами: аналог классического кафедрального журнала (в табличной форме) и более информативный – графический – вариант. Визуальное представление автоматически формируется в момент создания преподавателем урока и так же автоматически заполняется после выполнения заданий студентами. Данный блок

позволяет оперативно оценивать общую картину полученных студентами баллов и освободить преподавателя от просмотра каждого выполненного упражнения. Помимо среднего балла, за урок преподаватель получает детальную информацию о выполнении тех или иных заданий. Преподаватель может сам вмешаться в систему оценивания, дать свои комментарии и скорректировать траекторию обучения.

Выводы

Разрабатываемая в ТПУ интеллектуальная обучающая система является универсальным лингвистическим тренажером, позволяющим решить ряд актуальных методических и организационных задач. Ее концепция опирается на современные разработки в сфере технологий искусственного интеллекта и интернет-технологий, а практика внедрения в учебный процесс 1-го модуля тренажера «КЛИОС» «Система именного склонения» подтверждают ее функциональность, эффективность и удобство.

Список литературы

1. Брусиловский П., Миллер Ф. Адаптивные и интеллектуальные технологии для Сетевого обучения [Электронный ресурс]. – Режим доступа: <http://ifets.ieee.org/> (дата обращения 22.03.2013).
2. Горисев С.А. Интеллектуальный лингвопроцессорный комплекс для обучения русскому языку как иностранному / С.А. Горисев, А.В. Койнов, С.В. Лисинин, А.П. Савинов // Научно – методическая конференция «Уровневая подготовка специалистов: государственные и международные стандарты инженерного образования», 26-30 марта 2013 г. – Томск: Изд-во Томского политехнического университета, 2013. – С. 315-316.
3. Как работает Google переводчик [Электронный ресурс]. – Режим доступа: http://translate.google.com/about/intl/ru_ALL/ (дата обращения 17.03.2013).
4. Обзор основных элементов в Moodle [Электронный ресурс]. – Режим доступа: <http://moodle-center.ru/video/12> (дата обращения 05.04.2013).
5. Смирнова Н.В. Следящие интеллектуальные обучающие системы: состояние и перспективы / Н.В. Смирнова // Интеллектуальные системы управления. Под ред. Васильева С.Н. – М.: Изд-во «Машиностроение», 2010. – С. 434-446.
6. АBBYU Compreno [Электронный ресурс]. – Режим доступа: <http://www.abbyu.ru/science/technologies/business/compreno/> (дата обращения 07.04.2013)

7. Balackboard Review & Rating [Электронный ресурс]. – Режим доступа: <http://www.pcmag.com/article2/0,2817,17187,00.asp> (дата обращения 06.04.2013).
8. Hot Potatoes [Электронный ресурс]. – Режим доступа: [http:// web.uvic.ca/hrd/hotpot/](http://web.uvic.ca/hrd/hotpot/) (дата обращения 01.04.2013).
9. Pymorphy 2 [Электронный ресурс]. – Режим доступа: <http://pymorphy2.readthedocs.org/> (дата обращения 06.04.2013).

Рецензенты:

Гришин А.М., д.ф.-м.на., профессор, зав. кафедрой «Физической и вычислительной механики» Национального исследовательского Томского государственного университета, г. Томск.

Тузовский А.Ф., д.т.н., доцент, профессор кафедры «Оптимизации систем управления» института «Кибернетики» национального исследовательского Томского политехнического университета, г. Томск.