

УДК 007:378

К.А. МЕТЕШКИН, Л.А. ФЕДОРЧЕНКО

*Международный славянский университет, Харьков, Украина***ЗАДАЧА И ПРИНЦИПЫ ПОСТРОЕНИЯ ИНСТРУМЕНТАЛЬНЫХ СРЕДСТВ ЛИНГВОСЕМАНТИЧЕСКОЙ ПОДДЕРЖКИ ОБРАЗОВАТЕЛЬНЫХ ПРОЦЕССОВ**

В работе предложена постановка задачи и сформулированы принципы построения корпусов текстов учебного назначения для формирования профессиональных языков в вузах.

образовательный процесс, искусственный интеллект, корпус текстов, лингвoseмантическая поддержка, профессиональный язык

Информационная революция, по мнению многих ученых, является первопричиной интенсификации труда людей, в основе которого лежит интеллектуальная деятельность [1, 2]. Во многих научных публикациях и в частности [2 – 4], обосновывается объективная необходимость перехода к новым информационным технологиям во всех сферах человеческой деятельности, предусматривающим построение и эффективное использование систем с интегрированным интеллектом. Особенно ощутима эта необходимость в образовательной сфере, где за последнее десятилетие возросла интенсивность трудовой деятельности научно-педагогических работников вузов [5].

Возникает противоречие между непрерывно увеличивающимся объемом информации, которую необходимо перерабатывать преподавателю для подготовки к занятиям, и его фиксированной учебной нагрузкой. Обострение этого противоречия приводит к неустойчивости психофизиологического состояния преподавателей и отражается на качестве их преподавания.

Указанный факт порождает проблему повышения качества образовательных процессов в вузах в условиях интенсификации педагогической деятельности.

Можно утверждать, что эту проблему в настоящее время пытаются решить по нескольким направлениям. Во-первых, директивно, т.е. перераспреде-

ляя время самостоятельной работы студентов и время их учебы под руководством преподавателей в сторону увеличения самостоятельной учебы студентов. Во-вторых, унификацией и стандартизацией отдельных блоков (циклов) учебных дисциплин [6]. В-третьих, на основе применения в преподавательской деятельности информационных и коммуникационных образовательных технологий (дистанционного образования) [7]. В-четвертых, использованием в педагогической практике обучающих систем, построенных на основе методов искусственного интеллекта, где используемые языково-информационные процессы придают языку технологический статус [2, 3]. В-пятых, созданием инновационных образовательных технологий, в которых используется интегрированный интеллект, т.е. естественный интеллект преподавателей и модели их профессиональных знаний [8, 9].

Предложим еще одно направление, основанное на формировании некоторой совокупности текстов, имеющих учебно-методическую и научную направленность и их автоматическом (автоматизированном) лингвoseмантическом анализе.

Целью настоящей статьи является постановка задачи и формулировка принципов построения инструментальных средств, обеспечивающих лингвoseмантическую поддержку образовательных процессов в вузах.

Последние достижения в области корпусной лингвистики [9], а также результаты работ по созданию Британского национального корпуса (BNC; <http://sara.natcorp.ox.ac.uk>), национального корпуса русского языка (сайт <http://ruscorpora.ru>) и интегральной лексикографической системы «Словники України» [2] показывают возможность и большую перспективу создания корпусов профессиональных языков в вузах.

Под корпусом профессиональных языков (КПЯ) будем понимать взаимосвязанную совокупность корпусов текстов, касающуюся подготовки студентов по конкретной специальности вуза. Корпус текстов должен формироваться в электронном виде и состоять из учебно-методических материалов (текстов учебников, учебно-методических пособий, конспектов лекций, методических разработок, учебных программ, словарей и т.д.). Кроме того, в состав корпуса текста целесообразно помещать научную литературу, имеющую отношение к конкретной специальности, по которой формируется корпус профессионального языка. В качестве научного текста можно использовать тексты монографий, научных статей, авторефераты диссертаций, диссертации и т.д.

Для достижения поставленной цели воспользуемся дефинициями и результатами анализа структур дидактических языков, которые введены в работе [10].

Язык науки – это инструментальные средства познания действительности в различных предметных областях.

Дидактический язык науки – это язык, при помощи которого излагается в вузах учебный материал в устной или письменной форме.

Носители дидактического языка науки – научно-педагогические работники вуза, владеющие дидактическими основами и могущие излагать суть изучаемых вопросов той или иной науки (учебной дисциплины). Под дидактическими основами будем

понимать грамматику дидактического языка науки или, другими словами, методику преподавания учебных дисциплин.

Профессиональный язык – это язык, при помощи которого выпускник вуза может излагать сущность и способы решения специальных задач, заданных квалификационными характеристиками на выпускника по той или иной специальности вуза.

Носители профессионального языка – выпускники по той или иной специальности вуза.

Корпус профессионального языка должен отличаться от известных корпусов своей целевой функцией.

Если корпуса национальных языков предназначены для изучения и исследования языков как таковых, то корпуса профессиональных языков предназначены для лингвосемантической поддержки образовательных процессов в высших учебных заведениях по конкретным специальностям вуза.

Сформулируем основные принципы создания инструментальных средств лингвосемантической поддержки образовательных процессов, построенных на основе формирования корпусов профессиональных языков.

Принцип адекватности корпусов текста дидактическим языкам конкретной специальности.

Данный принцип должен обеспечивать формирование некоторой совокупности текстовой информации по конкретной учебной дисциплине. Например, текстов учебной программы, тематического плана, учебника, учебного пособия, методических разработок, список курсовых работ и т.д. Такую совокупность текстов будем называть корпусом дидактического языка (КДЯ).

Принцип системологической целостности корпуса профессионального языка.

Корпус профессионального языка должен обладать свойствами системы, у которой корпуса дидактических языков имеют между собой логические связи, соответствующие связям структурно-логи-

ческой схемы подготовки студентов по конкретной специальности, а также отображать семантическую сеть ее терминов.

Принцип лексикографического обеспечения корпуса профессионального языка.

Лексикографическое обеспечение является одним из основных принципов построения корпуса профессионального языка. Электронные словари должны составлять специальный корпус лексикографических произведений (КЛП), который должен пополняться терминами и определениями, введенными в научных текстах.

Принцип открытости корпуса профессионального языка.

Корпус профессионального языка должен иметь возможность пополнения за счет расширения (добавления) корпусов текстов в корпуса дидактических языков.

Принцип многофункциональности корпуса профессионального языка.

Корпус профессионального языка должен не только обеспечивать лингвосемантическую поддержку реализации педагогических решений, но и обеспечивать решение научно-исследовательских и других задач. Например, задач контроля за обновлением текстов дидактических языков; задач контроля непротиворечивости и нетавтологичности используемых терминов.

Принцип оперативности использования новых научных разработок (результатов защищенных диссертаций) в образовательных процессах.

Реализация данного принципа должна обеспечивать носителей дидактических языков возможностью оперативной корректировки сути излагаемого материала с учетом последних достижений науки в заданной предметной области.

Принцип многоязычности корпуса профессионального языка.

Этот принцип должен обеспечивать создание параллельных корпусов дидактических языков, текст

которых представлен на украинском, английском, немецком и других языках. При этом основной корпус формируется на русском или украинском языке.

Принцип интеллектуализации взаимодействия пользователей с корпусом профессионального языка.

Взаимодействие пользователей с корпусом профессионального языка должен обеспечивать интеллектуальный интерфейс, который позволял бы осуществлять как автоматическую (автоматизированную) разметку корпусов текстов, так и осуществлять быстрый поиск текстовой информации заданной конфигурации и с заданными параметрами.

Покажем на рис. 1 структурную схему инструментальных средств лингвосемантической поддержки образовательных процессов в вузе, ядром которой является корпус профессиональных языков.

На рис. 1 обозначено: КДЯ – ГП – корпус дидактических языков, относящийся к гуманитарной подготовке студентов; КДЯ – ФП – корпус дидактических языков, относящийся к фундаментальной подготовке студентов; КДЯ – ПП – корпус дидактических языков, относящийся к профессиональной подготовке студентов.

Из рис. 1 видно, что корпус профессионального языка имеет сложную структуру, элементы которой связаны между собой отношением включения. Кроме того, для решения лингвосемантических задач поддержки образовательных процессов в вузе необходимо задать некоторое множество отношений между элементами корпуса профессионального языка, между элементами множества корпусов дидактических языков, между элементами множества корпусов текстов, а также элементами корпусов лексикографических произведений и всеми остальными элементами структуры корпуса профессиональных языков.

Выделим несколько уровней разметки текста. На первом уровне разметки, назовем его метатекстовой разметкой, необходимо приписать текстам атрибу-

ты, которые бы позволили сформировать корпуса дидактических языков. Они, как видно из рис. 1, состоят из трех типов корпусов текстов: корпуса

учебно-методического текста; корпуса научных текстов; корпуса текстов лексикографических произведений.

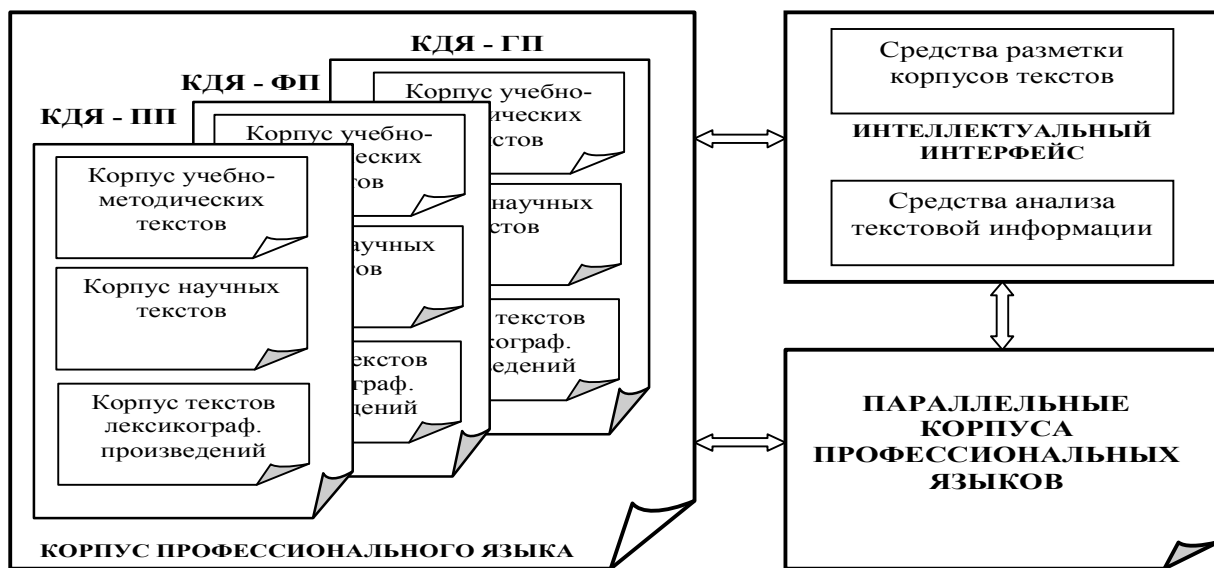


Рис. 1. Обобщенная структурная схема инструментальных средств лингвосемантической поддержки образовательных процессов

Очевидно, метатекстовую разметку необходимо производить на основе экспертных методов. Один из таких методов приведен в работе [11], где разработан метод классификации словарно-справочных средств по признаку их принадлежности к тому или иному блоку учебных дисциплин (корпусу дидактических языков).

Одной из важных процедур метатекстовой разметки является выявление логических связей и их силы между текстами конкретных корпусов дидактических языков. Другими словами, необходимо ответить на следующие вопросы. Является ли семантика некоторого учебно-методического текста обеспечивающей или обеспечиваемой другие учебно-методические тексты? Используются ли в рассматриваемом учебно-методическом тексте термины и их дефиниции, которые уже использовались в текстах других дидактических языков, или они являются новыми для студентов на данном промежутке времени?

Для выявления этих связей создана специальная программа [12], которая позволяет на основе оценок

группы носителей дидактических языков (научно-педагогических работников вуза), имеющих непосредственное отношение к созданию корпуса профессионального языка создать семантическую сеть конкретных КДЯ. Прототипом такой сети является структурно-логическая схема подготовки студентов по конкретной специальности вуза.

Второй уровень разметки касается непосредственно конкретного корпуса дидактического языка. Здесь каждому тексту, помещаемому в корпус профессионального языка, задаются атрибуты, которые относят его к учебно-методическому, научному тексту или тексту лексикографических произведений. Формализация учебно-методических текстов приведена в работе [4], а конкретное представление корпусами текстов учебной программы [13] и словарей различного назначения приведено в работах [14].

Третий уровень разметки будем осуществлять с учетом рекомендаций, приведенных на сайте (<http://ruscorpora.ru>). Особенностью разметки этого уровня является выделение в тексте ключевых терминов и присвоение им специальных помет. Эту

процедуру можно осуществить либо автоматически, задавая правила выделения отдельных терминов, либо вручную формировать список ключевых терминов, как это делается при оформлении авторефератов диссертаций.

Дефинитный уровень разметки будем характеризовать как присвоение паре «термин-определение» специальных помет, указывающих на то, что по тексту вводятся определения. Здесь может быть также два варианта. В первом варианте текст содержит слово «Определение», и оно является признаком того, что в тексте вводится новое определение. Во втором варианте из текста необходимо выделить определения терминов, которые могут содержать иерархию родовидовых отношений с другими терминами, определенными в тексте. Для выделения таких дефинитных деревьев необходимо разработать специальную процедуру, которая позволяла бы на основе исходного текста генерировать цепочки помет для построения семантической сети терминов предметной области.

Создание системы лингвосемантической поддержки образовательных процессов в вузе потребует разработки не только процедур разметки текстов, но и специального языка «запросов и ответов». Кроме того, потребует использования уже существующих методов и процедур анализа и синтеза текстовой информации, а также разработку новых методов и средств, обеспечивающих более глубокий семантический анализ текстов, в частности, методов выделения из текста различных видов терминов и их определений. Для этих целей могут быть использованы результаты исследований, приведенные в монографии [15], где описаны грамматические и семантические особенности построения текстов определений терминов.

Выделение из фрагментов текста текстовых маркеров, наименований терминов, направления и силы связи между ними позволят обеспечить автоматическое (автоматизированное) построение терминологических

сетей учебных дисциплин и специальности.

Таким образом, на основе вышесказанного сделаем *следующие выводы*.

На основе последних достижений в области корпусной лингвистики и лексикографии предложена новая задача создания системы лингвосемантической поддержки образовательных процессов в вузах, а также сформулированы основные принципы ее построения.

Разработана обобщенная структурная схема системы лингвосемантической поддержки образовательных процессов, основным ядром которой является корпус профессионального языка. Его тексты, размеченные специальным образом, являются исходными данными для решения таких задач, как создание и обновление учебного материала конкретных учебных дисциплин, организация и осуществление контроля знаний обучающихся с учетом смежных учебных дисциплин, осуществление сжатия, с различной степенью, учебного материала и автоматизированное построение семантической сети терминов в учебном материале, исследование закономерностей «старения учебного материала», исследование закономерностей изучения студентами иностранных языков и многие другие задачи, связанные с использованием языка в образовательных процессах. Корпуса профессиональных языков могут служить исходным материалом для разработки терминологических словарей и терминологических стандартов для различных отраслей знаний.

Создание параллельных корпусов профессиональных языков в системе лингвосемантической поддержки образовательных процессов позволит организовать исследования, связанные углубленным сравнительным анализом языков.

При создании и непрерывном сопровождении корпусов профессиональных языков в разных вузах возникает возможность их сравнения и объединения, что повышает их репрезентативность.

Совершенствование лингвистического обеспечения образовательных систем высшей школы, на наш взгляд, позволит в какой-то мере разрешить противоречия, которые указаны во вводной части настоящей статьи.

Литература

1. Формирование общества, основанного на знаниях. Новые задачи высшей школы. – М.: Изд-во «Весь Мир», 2003. – 232 с.
2. Широков В.А. Феноменология лексикографических систем. – К.: Наук. думка, 2004. – 327 с.
3. Петрушин В.А. Экспертно-обучающие системы. – К.: Наук. думка, 1992. – 196 с.
4. Белова Л.А., Метешкин К.А., Уваров О.В. Логико-математические основы управления учебными процессами вузов. – Х.: Восточно-региональный центр гуманитарно-образовательных инициатив, 2001. – 272 с.
5. Чоговадзе Г.Г. Информация: информация, общество, человек. – М.: ООО Дата+, 2003. – 320 с.
6. Дубичинский В.В., Метешкин К.А., Федорченко Л.А. Задача терминологической стандартизации образовательных процессов высшей школы и пути ее решения // Проблемы инженерно-педагогической освіти. Зб. наук. праць. – Х.: УІПА. – 2004. – № 7. – С. 94 – 100.
7. Кухаренко В.М., Рибалко О.В., Сиротенко Н.Г. Дистанційне навчання: Умови застосування. Дистанційний курс: Навчальний посібник / За ред. В.М. Кухаренка. – Х.: НТУ «ХПІ», «Торсінг», 2002. – 320 с.
8. Метешкин К.А. Кибернетическая педагогика: теоретические основы управления образованием на базе интегрированного интеллекта. Монография. – Х.: Межд. Славянский ун-т, 2004. – 400 с.
9. Баранов А.Н. Введение в прикладную лингвистику: учебное пособие. – М.: Эдиториал УРСС, 2001. – 360 с.
10. Метешкин К.А. Анализ структуры дидактических языков наук // Проблемы построения лингвистического обеспечения образовательных систем высшей школы. Сборник научных статей. Международный Славянский университет. – Х.: Экограф. – 2005. – С. 39 – 49.
11. Метешкин К.А., Федорченко Л.А., Жадан А.И. Задача и метод экспертизы лексикографических произведений // Вестник Харьковского национального автомобильно-дорожного университета. Сборник научных трудов. – Х.: ХНАДУ. – 2004. – Вып. 25. – С. 13 – 17.
12. Метешкин К.А., Довнар А.И., Раковская Башмакова О.С. Метод экспериментального исследования междисциплинарных связей учебных дисциплин в высшем учебном заведении // Системы обработки информации. – Х.: ХВУ. – 2004. – Вып. 11 (39). – С. 126 – 132.
13. Метешкин К.А., Федорченко Л.А., Кобзистая Н.А. Учебная программа как объект исследования корпусной лингвистики // Радиоелектронні і комп'ютерні системи. – Х.: Нац. аерокосм. ун-т ім. М.Є. Жуковського «ХАІ». – 2004. – №3 (7). – С. 59 – 62.
14. Метешкин К.А., Самойлов А.Н., Федорченко Л.А. Лексикографическое произведение как объект исследования корпусной лингвистики // Открытые информационные и компьютерные интегрированные технологии. – Х.: НАКУ «ХАИ». – 2004. – Вып. 22. – С. 166 – 174.
15. Шелов С.Д. Термин. Терминологичность. Терминологические определения (филологические исследования). – С.-Пб.: Филологический факультет СПбГУ, 2003. – 280 с.

Поступила в редакцию 1.04.05

Рецензент: д-р техн. наук, проф. О.Е. Федорович, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.