

НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК КЫРГЫЗСКОЙ РЕСПУБЛИКИ
Институт физико-технических проблем и материаловедения им. Ж. Жеенбаева

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ КЫРГЫЗСКОЙ РЕСПУБЛИКИ
Кыргызский государственный технический университет им. И. Раззакова
Кыргызско-Российский Славянский университет им. Б. Ельцина

Диссертационный совет Д.05.11.034

На правах рукописи

УДК 004.42:811

Долматова Полина Сергеевна

**АВТОМАТИЗИРОВАННАЯ СИСТЕМА ИНТЕРАКТИВНОГО
ОПЕРАЦИОННОГО ПРЕДСТАВЛЕНИЯ ПОНЯТИЙ ЕСТЕСТВЕННЫХ
ЯЗЫКОВ**

Специальность **05.13.06** – Автоматизация и управление технологическими процессами и производствами (образование)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Бишкек-2012

Работа выполнена в Кыргызско-Российском Славянском университете им. Б. Ельцина.

Научный руководитель: доктор физ.-мат. наук, профессор, член-корр.
НАН КР **Панков Павел Сергеевич**

Официальные оппоненты: академик ИА КР, доктор технических наук,
профессор **Муслимов Аннас Поясович**

кандидат технических наук, доцент
Осмонов Мелис Сыдыкбекович

Ведущая организация: Институт интеграции международных образовательных программ, факультет компьютерных технологий и интернет КНУ им. Ж. Баласагына, адрес 720033, г. Бишкек, ул. Фрунзе 547.

Защита состоится «25» января 2013 года в 16-00 часов на заседании Межведомственного диссертационного совета Д.05.11.034 при Институте физико-технических проблем и материаловедения им. Ж. Жеенбаева НАН КР, Кыргызском государственном техническом университете им. И. Раззакова, Кыргызско-Российском Славянском университете им. Б. Ельцина, по адресу: 720071, г. Бишкек, пр. Чуй, 265-а, центральный корпус НАН КР.

С диссертацией можно ознакомиться в Центральной научной библиотеке Национальной академии наук Кыргызской Республики.

Автореферат разослан «21» декабря 2012 г.

Ученый секретарь Межведомственного диссертационного совета Д 05.11.034
кандидат физико-математических наук



В. В. Алиферов

ОБЩАЯ ХАРАКТЕРИСТИКА

Актуальность темы диссертации. В настоящее время в мировом сообществе интенсивно развиваются процессы глобальной информатизации всех сфер общественной жизни. Во всех развитых странах и во многих развивающихся странах идут процессы информатизации образования. Разрабатываются пути повышения эффективности общего образования, вкладываются большие средства в разработку и внедрение новых информационных технологий в образовательную деятельность.

Общепризнанно, что современная система образования вступила с появлением компьютеров в новую фазу своего развития, благодаря интенсивному освоению возможностей новых информационных технологий. В этом контексте актуальной задачей становится совершенствование методик обучения применительно к новым образовательным условиям. Усилия многих теоретиков и практиков образования сегодня сосредоточены в области использования компьютерных технологий для обучения, в том числе естественным языкам.

В целом, анализ научных исследований показывает, что, несмотря на значительное продвижение, наметившееся в последнее время в реализации различных компьютерных технологий обучения, в том числе и обучения естественным языкам, потенциал информационных технологий реализуется в учебно-воспитательном процессе еще недостаточно из-за отсутствия мультимедийных учебных комплексов, алгоритмов их разработки и применения, методических рекомендаций по их эффективному применению.

Кроме того, ни в одной стране, в том числе и в Кыргызской Республике, не была разработана методика комплексного компьютерного экзамена по государственному языку.

Таким образом, теоретическая и практическая значимость, недостаточная разработанность методологии и методики компьютерных технологий обучения естественным языкам при достаточной аппаратной базе явились причиной выбора проблемы исследования, которая заключается в поиске и реализации основ для построения автоматизированной системы, которая позволяла бы создавать независимые обучающие и комплексные экзаменуемые приложения по естественным языкам, основанные на интерактивном операционном представлении понятий естественных языков, экспериментальной апробации и внедрении данной автоматизированной системы в процесс обучения естественным языкам. Данная проблема обусловила выбор темы диссертационного исследования: “Автоматизированная система интерактивного операционного представления понятий естественных языков”.

Связь темы диссертации с крупными научными программами, основными научно-исследовательскими работами, проводимыми научными учреждениями. Данная работа была выполнена в рамках следующих проектов Национальной Академии наук Кыргызской Республики:

«Компьютеризация грамматики кыргызского языка и интерактивное представление основных понятий кыргызского языка» (2007 г.) и «Разработка программного обеспечения для официального экзамена и электронного учебника по кыргызскому языку» (2008 – 2009 г.).

Цель и задачи исследования. Цель исследования состоит в разработке автоматизированной системы, предназначенной для интерактивного операционного представления понятий естественных языков и создания независимых обучающих и комплексных экзаменуемых приложений по естественным языкам.

Объект исследования – процесс обучения естественным языкам с помощью компьютера.

Предмет исследования – интерактивное операционное представление понятий естественных языков.

Для достижения цели исследования были поставлены следующие задачи:

1. Разработать теоретические основы интерактивного представления понятий естественных языков; с этой целью использовать операционный принцип – обратная связь осуществляется при помощи действий пользователя по методике Drag-and-Drop; усовершенствовать эту методику для расширения набора операций.

2. Построить алгоритмический язык, позволяющий осуществлять операционное представление понятий естественных языков.

3. Разработать автоматизированную систему, позволяющую осуществлять операционное представление понятий естественных языков и создавать независимые обучающие и комплексные экзаменуемые приложения по естественным языкам.

4. Провести экспериментальную апробацию созданной автоматизированной системы и разработать рекомендации по её использованию.

5. Реализовать первую версию комплексного экзамена и основу для разработки независимого электронного учебника по кыргызскому языку.

Научная новизна исследования состоит в:

- разработке новых определений: математическая модель понятия естественного языка, минимальная математическая модель понятия естественного языка, первичное понятие естественного языка, предшествующее понятие естественного языка, компьютерная модель понятия естественного языка;

- формулировке математических моделей основных понятий естественных языков;

- разработке математических и компьютерных методов представления объектов и преобразований, необходимых для интерактивного представления понятий естественных языков;

- разработке способа автоматизации операционного представления понятий естественных языков;

- создании и реализации алгоритмического языка, позволяющего осуществлять операционное представление понятий естественных языков.

Практическая значимость исследования заключается в следующем:

- разработанная и апробированная автоматизированная система позволяет специалистам-филологам непосредственно участвовать в создании авторского учебника или экзамена по языку;

- полученные результаты могут быть использованы для создания аналогичных программных средств для изучения и проверки знаний по естественным языкам.

Основные положения диссертации, выносимые на защиту.

1. Теоретические основы интерактивного представления понятий естественных языков с использованием операционного принципа.
2. Алгоритмический язык, позволяющий осуществлять операционное представление понятий естественных языков.
3. Автоматизированная система интерактивного операционного представления понятий естественных языков и создания независимых обучающих и комплексных экзаменуемых приложений по естественным языкам.

Личный вклад соискателя. Результаты и положения, выносимые на защиту, получены лично соискателем.

Апробации результатов диссертации.

Материалы работы докладывались на Международной конференции «Проблемы науки и образования в условиях глобализации», КНУ им. Ж. Баласагына (26 – 27 апреля 2008 г.), конференции «Проблемы подготовки специалистов в области информационных технологий», КНУ им. Ж. Баласагына (17 – 18 ноября 2008 г.), на межвузовской студенческой конференции, МУК (15 апреля 2008 г.), Международной юбилейной научной конференции, посвященной 15-летию КРСУ «Актуальные проблемы теории управления, топологии и операторных уравнений», КРСУ им. Б. Ельцина (22 – 25 сентября 2008 г.), расширенном научном семинаре кафедры информационных и вычислительных технологий и кафедры информатики Кыргызско-Российского Славянского университета (2 декабря 2011 г.), Международной научно-практической конференции «Современный университет: стратегия развития в меняющемся мире» в ИИМОП КНУ им. Ж. Баласагына (20 апреля 2012 г.), Академических Инновационных Инициативах-2012 «Наука и Национальный интерес» НАН КР (29 мая 2012 г.); демонстрировались на выставке научных разработок НАН КР «Инновации-2012» (9 ноября 2012 г.).

Полнота отражения результатов диссертации в публикациях.

По теме работы опубликовано 12 статей. В совместных публикациях [1] – [6] научному руководителю принадлежит постановка задачи, автору – полученные результаты.

Структура и объём диссертации. Диссертационная работа состоит из введения, 4 глав, выводов, списка использованных источников, приложений. Работа содержит 199 страниц компьютерного текста, 54 иллюстрации, 15 таблиц, 3 приложения и список использованных источников из 104 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована важность и актуальность темы диссертации, сформулированы цели диссертационной работы и решаемые задачи, определяется научная новизна работы и её практическая значимость, а также представлены апробация и полнота отражения полученных результатов на конференциях и в публикациях.

В первой главе даётся обзор публикаций по рассматриваемой теме. Указаны полученные другими авторами результаты в области исследования и моделирования понятий языков, отмечаются как методы и разработки, используемые и развиваемые в данной работе, так и отсутствие каких-либо аналогов результатов данной работы для доказательства её новизны.

Ранее были реализованы приложения компьютерной геометрии к преобразованию объектов на экране пользователем с помощью клавиатуры и мыши, однако подобные действия не использовались для тестирования знаний и изучения естественных языков.

Существуют программные средства для составления компьютерных тестов, позволяющие использовать вопросы закрытого типа и случайный выбор порядка ответов в вопросе и вопросов в тесте, но программных средств, которые бы давали возможность вводить ответ при помощи непрерывных преобразований, ранее не было.

Существуют программные средства для изучения естественных языков на компьютере, но в них либо присутствуют инструкции или перевод на известный пользователю язык, либо при отсутствии таких инструкций и перевода имеется лишь фиксированный набор заданий. Методика Total Physical Response (TPR) достаточно широко использовалась для изучения языков без помощи специальных компьютерных программ. В программных продуктах для изучения языков с использованием этой методики, авторы ограничивались лишь одним видом заданий – из фиксированного множества объектов выбрать и захватить курсором мыши один и перетащить на главную картинку. Кроме того, в таких программных средствах не было предусмотрено формирование случайных заданий.

В рамках моделирования естественных языков Л. С. Выготский, Л. С. Сахаров, Terry Winograd разрабатывали представление одних понятий через другие и связи между понятиями, В. А. Тузовым предложена функциональная

модель языка. В Кыргызской Республике интерактивным представлением естественных языков занимались П. С. Панков, М. Жураев, Б. Ж. Баячорова.

П. С. Панковым было предложено следующее

Определение 1. Пусть дано какое-либо "Понятие" ("Слово" из языка).

Если алгоритм, действуя на компьютере:

- демонстрирует человеку достаточно большое количество ситуаций;
- в каждой ситуации даёт команду с этим "Словом";
- принимает действия человека и наглядно демонстрирует их результаты;
- определяет, соответствуют ли эти результаты данной команде,

то такой алгоритм назван компьютерным интерактивным представлением "Понятия".

В связи с тем, что данная работа выполнялась в рамках проекта «Разработка электронного учебника и официального экзамена по кыргызскому языку», был также проведен обзор публикаций по формализации и алгоритмизации кыргызского языка. Этими вопросами занимались К. Тыныстанов, Б. О. Орузбаева, Т. Садыков, Р. А. Чодоев и др.

Таким образом, несмотря на большое разнообразие программных продуктов для изучения и проверки знаний по естественным языкам, ранее не существовало единой программной реализации, объединяющей следующие идеи:

- изучения естественного языка без использования языка-посредника;
- использования случайного выбора при формировании заданий;
- методики TPR;
- использования преобразований объектов на экране в соответствии с алгоритмами компьютерной геометрии в графических заданиях;
- доступности для специалистов-филологов, которые сами создают обучающий и контролирующий материал.

Во второй главе рассмотрены теоретические основы, а также математические и компьютерные модели понятий естественных языков, на основе которых строится автоматизированная система. Описан неформальный алгоритмический язык для независимого интерактивного представления понятий.

В качестве математического аппарата был выбран язык теории множеств, поскольку понятия естественного языка могут быть рассмотрены по нашему методу, как действия с объектами, математически представляемыми в виде связанных множеств.

В качестве реализации **Определения 1** предложено

Определение 2. Математическая модель понятия естественного языка (ММП) – это:

- список множеств, некоторые из которых являются подмножествами других множеств из списка, с указанием наборов множеств для случайного выбора;
- список возможных движений и других преобразований этих множеств;

- списки допустимых и недопустимых отношений между множествами при их преобразованиях;
- список необходимых отношений между множествами (непустое пересечение, вложение) во временной последовательности;
- список заданий, содержащих данное понятие, дающих его достаточно полное и адекватное представление.

Также предложено

Определение 3. Компьютерная модель понятия естественного языка (КМП) – модель, построенная на основе ММП с графическим представлением множеств.

Далее рассмотрен каждый из вышеприведенных списков в рамках **Определения 3.**

Пусть задан прямоугольник D с обычной евклидовой метрикой (экран дисплея). От реального экрана прямоугольник D отличается тем, что он непрерывен, но современные дисплеи с большой разрешающей способностью также создают ощущение непрерывности. Для простоты ограничимся двуцветным (черно-белым) изображением и введём множество 2^D всевозможных подмножеств D . Обозначим через $J \subset 2^D$ некоторое множество. Будем считать, что J содержит только замкнутые множества и потребуем, чтобы множество J было связно, т. е. чтобы любую реализацию можно было получить из любой другой реализации непрерывным переходом внутри J . Тогда множество J будем называть множеством реализаций некоторого объекта на экране.

Рассмотрены возможные преобразования реализаций объекта на экране. Анализ показывает, что различным понятиям соответствуют существенно различные допустимые преобразования. Например, существительное "книга" допускает преобразования растяжения (сжатия), подобия, сдвига и поворота, "дерево" – растяжения (сжатия), подобия и сдвига, "1 см" – сдвига и поворота, "шар" – подобия и сдвига.

Также, различным глаголам соответствуют различные математические преобразования. Так, например, глагол *двигать(ся)* может быть рассмотрен как аналог любого движения (параллельного переноса, поворота, симметрии), а также любых непрерывных преобразований в топологических пространствах. *Толкать* и *бросать* соответствуют собственным движениям; *крутить(ся)* соответствует повороту без параллельного переноса; *удалять(ся)*, *приближать(ся)* соответствуют только параллельному переносу. Такой глагол, как *растягивать(ся)*, не применимый к рассматриваемым нами твердым телам, описывается топологией (сохраняет топологические инварианты) и его можно также поставить в соответствие аффинным преобразованиям. Глагол *завязывать* описывается топологией в трёхмерном пространстве. Глагол *резать* оказывает такие действия, что топологические свойства исходного предмета не сохраняются.

Каждому глаголу сопоставлено пространство наименьшей размерности, в котором математически описываются инварианты глагола. «Одномерными» являются следующие глаголы: *удалять(ся)*, *приблизжать(ся)*, *падать*, *поднимать(ся)*, *опускать(ся)*, *колебаться*, *резать*, *соединять*, *растягивать(ся)*, *тащить*. Для описания смысла этих глаголов достаточно продемонстрировать пользователю соответствующий процесс на прямой. Значение таких глаголов, как *крутить(ся)*, *поворачивать(ся)*, *смотреть*, *обходить*, *гнуть(ся)*, *выпрямлять(ся)*, *окружать*, раскрывается только в двумерном пространстве (на плоскости). «Трёхмерными» являются, например, глаголы *разглаживать*, *завязывать*, *развязывать*, *обволакивать*. Но для некоторых глаголов не требуется непрерывного пространства, а достаточно дискретной модели. Примерами таких глаголов являются понятия *искать*, *интересоваться*, *выбирать*.

Далее, рассмотрена задача о достаточно гибких геометрических преобразованиях на дисплее, возникающая в результате реализации описанных преобразований в рамках **Определения 3**. Рассмотрен вопрос об осуществлении пользователем движения и поворотов графических объектов на экране дисплея.

В множестве J введена хаусдорфова метрика, которая больше всего соответствует интуитивному представлению “близости” изображений.

Обозначим через $U(t)$ действие пользователя и через $Z(t)$ – состояние процесса (элемент множества J) в момент времени t .

Если множество J связное и гладкое, то возникает условное уравнение, дифференциальное по $Z(t)$ и локальное по $U(t)$ в том смысле, что правая часть зависит только от значений U в точке t и сколь угодно близких к ней точках:

$$Z'(t) = F(Z(t), U(t + \Delta t); \Delta t \rightarrow 0), \quad (1)$$

где $F(Z, U)$ – заданная оператор-функция.

Требования к $U(t)$ и $F(Z, U)$ поставим формально: они должны обеспечивать перевод из любого элемента Z_0 множества J в любой другой элемент Z_1 , т. е. должна иметь решение краевая задача

$$Z(0) = Z_0, Z(T) = Z_1 \quad (2)$$

для некоторого $T > 0$;

и неформально: произвольным образом выбирая $U(t)$ и наблюдая за $Z(t)$, пользователь должен быстро догадаться о свойствах функции $F(Z, U)$ и способе решения задачи (1) – (2); число T должно быть «не очень большим».

Для решения задачи (1) – (2) построена следующая система векторного дифференциального уравнения со скалярным произведением

$$P'(t) = \langle U(t) - P(t), U'(t) \rangle \|U(t) - P(t)\|^{-2} (U(t) - P(t)) \quad (3)$$

и алгебраического равенства (угол измеряется в радианах)

$$A(t) = \arg(U(t) - P(t)), \quad (4)$$

которое при заданном начальном значении $A(0)$ эквивалентно дифференциальному равенству

$$A'(t) = \langle Rot(U(t) - P(t), \pi/2), U'(t) - P'(t) \rangle \|U(t) - P(t)\|^{-2}, \quad (5)$$

где $U(t) = (u(t), v(t))$ – координаты курсора, которые задаёт пользователь при помощи компьютерной мыши в момент времени t ; $P(t) = (x(t), y(t))$ – перемещение одной из точек объекта; $A(t)$ – угол поворота объекта; $Rot(V, \beta)$ – оператор поворота вектора V на угол β .

Предложено выделить небольшой круг $S(t)$ (полностью находящийся внутри объекта) с центром в $P(t)$ и радиусом r с тем условием, чтобы при $U(t) = P(t)$ происходило только параллельное перемещение объекта, а при $U(t) \in S(t) \setminus P(t)$ осуществлялось бы движение, промежуточное между параллельным перемещением объекта и построением траектории. Получена следующая система из векторного дифференциального уравнения и дифференциального равенства:

$$P'(t) = \begin{cases} U'(t)(1 - \|U(t) - P(t)\|/r) + \langle U(t) - P(t), U'(t) \rangle \|U(t) - P(t)\|^{-1} (U(t) - P(t)) / r & \text{при } \|U(t) - P(t)\| < r, \\ \langle U(t) - P(t), U'(t) \rangle \|U(t) - P(t)\|^{-2} (U(t) - P(t)) & \text{при } \|U(t) - P(t)\| \geq r. \end{cases} \quad (6)$$

$$A'(t) = \begin{cases} \langle Rot(U(t) - P(t), \pi/2), U'(t) - P'(t) \rangle \|U(t) - P(t)\|^{-1} / r & \text{при } \|U(t) - P(t)\| < r, \\ \langle Rot(U(t) - P(t), \pi/2), U'(t) - P'(t) \rangle \|U(t) - P(t)\|^{-2} & \text{при } \|U(t) - P(t)\| \geq r. \end{cases}$$

Фактически экран дисплея дискретен, поэтому построена конечно-разностная аппроксимация системы (6).

Определение 2 включает три возможных вида отношений между множествами: допустимые, недопустимые и необходимые, взятые в их временной последовательности. Математически все отношения представляют собой операции с множествами, объединённые логическими операциями.

В главе 2 также приведены математические модели некоторых понятий кыргызского языка, построенные на основе **Определения 2**.

Для построения системы независимого изучения понятий языка в рамках **Определения 1** предложено

Определение 4. Первичное понятие (глагол или существительное) – это понятие, которое может быть понято пользователем без знания каких-либо других понятий.

Далее, на основе первичных понятий вводятся новые понятия.

Определение 5. Предшествующее понятие – понятие, необходимое для построения ММП данного понятия; понятие, которое должно быть изучено пользователем раньше, чем данное понятие.

На основе **Определений 4** и **5** формируется иерархия понятий, в соответствии с которой изучается естественный язык.

Также рассмотрены способы введения новых понятий:

- «естественность»: в данных условиях возможно лишь одно правильное (естественное) действие или последовательность действий;

- «подобие»: два объекта имеют аналогичное свойство и новое понятие вводится как наименование этого свойства;

- «различие»: появляется новый объект вместе с новым понятием как альтернатива ранее появлявшимся объектам и понятиям;

- «сравнение»: пользователю демонстрируется некоторый объект (или несколько объектов) с одним (или несколькими) характерным свойством, которое и служит критерием выбора объектов.

Далее, в данной главе на основе *Определения 1* построен неформальный алгоритмический язык для независимого компьютерного представления понятий, содержащий: формализованное подмножество естественного языка; утверждения, описывающие окружающую среду; утверждения, описывающие возможности пользователя; утверждения, описывающие возможности объектов; утверждения, описывающие условия. На основании *Определения 2* понятие в данном языке состоит из: списка объектов; списка предшествующих понятий; описания начальной окружающей среды (если это необходимо); команды, написанной в подмножестве естественного языка; последовательности условий (во временном порядке) с операциями И, ИЛИ, XOR, НЕ и действиями. Если все условия выполнены, тогда выдаётся сообщение Да (понятие было понято). Если какое-нибудь условие нарушено, то выдаётся сообщение Нет.

Таким образом, для представления понятий естественных языков предложено использовать математические модели этих понятий, определить множества, пространства и преобразования, которые им соответствуют, установить отношения между элементами множеств. Далее это реализуется на компьютере в виде соответствующей компьютерной модели понятия. Для этого множества представляются графически, а преобразования осуществляются с помощью манипулятора мыши. На основе математических моделей понятий естественных языков строится неформальный язык для независимого интерактивного представления естественных языков, который может быть реализован в виде некоторого алгоритмического языка.

Построение математических моделей понятий естественных языков с дальнейшей реализацией их в виде компьютерных моделей понятий было взято за основу при разработке автоматизированной системы Language Education System для создания независимых обучающих и комплексных экзаменуемых приложений по естественным языкам, а также алгоритмического языка программирования TaskLang, являющегося реализацией описанного выше неформального языка.

В третьей главе описан язык программирования TaskLang, созданный для написания обучающих и проверочных заданий по естественным языкам в специальном редакторе. На этом языке пишутся программы – набор инструкций языка TaskLang, который случайным образом формирует обучающее или проверочное задание.

Программирование на языке TaskLang состоит из: определения фиксированных объектов, формирования случайных объектов, составления случайных заданий. Программирование на языке TaskLang осуществляется в Language Education System – интерпретаторе языка TaskLang.

В языке TaskLang предусмотрена генерация случайных заданий, которая подразумевает, что написанное составителем задание каждый раз при очередном своем запуске будет выглядеть как новое задание, подобное, но не идентичное заданию при предыдущем запуске. В связи с этим были разработаны структура и элементы языка.

Объекты бывают двух типов – фиксированные и случайные (множество однородных фиксированных объектов). Фиксированный объект – это строка, которая может идентифицировать различные типы информации: текстовую, графическую и звуковую. Если использован случайный объект, то вместо него из соответствующего множества подставляется (случайным образом) фиксированный объект. Если случайный объект используется в задании несколько раз, то он принимает одно и то же значение. Можно запросить второе значение этого же объекта, и тогда он принимает отличное от первого значение, но оно остается постоянным. Если задание выполнить второй раз, то значения случайных объектов изменятся в соответствии с этим правилом.

Определение фиксированного объекта включает: имя объекта (object-name); класс, написанный на языке программирования C#, способный отвечать на внешние события и рисовать экземпляры данного класса (object-class); текстовое слово или список его грамматических форм (object-word); значения атрибутов.

Определение случайного объекта включает: имя объекта; список фиксированных объектов.

Случайное текстовое задание формируется как последовательность: строковых констант (включая слова, пробелы и знаки препинания); имен фиксированных объектов; имен случайных объектов (рассматриваемых в качестве ссылок на object-words) с определениями необходимых грамматических форм object-words; операторов языка TaskLang.

Составление графического задания включает в себя: расположение случайных объектов в графическом окне; задание числа слоёв для случайных объектов; задание направления медленного движения для некоторых случайных объектов; задание последовательности верных действий как правильного ответа на задание и набора неверных действий.

Приведем БНФ языка TaskLang.

ParseLanguage = { {TextCharacter}, Brackets, {TextCharacter} }

TextCharacter = (* Any UTF character except [=|#*~/ *)

Brackets = '[, Comment | FastChoise | KGZALG | MATH | DECLARE | GETPROPERTY | SETPROPERTY | DRAW | ANSWER | PLAYSOUND | DELAY

| EVENT | INCORRECTEVENT | GRASP | UNGRASP | ExtendedObject | Assignment | RandomAssignment, ']'

Comment = ParseLanguage

FastChoice = ParseLanguage, '|', ParseLanguage, {'|', ParseLanguage}

KGZALG = 'KGZALG*', ParseLanguage, '|', ParseLanguage

MATH = 'MATH*', NumericExpression

DECLARE = 'DECLARE*', ParseLanguage, '#', ParseLanguage, {'|', ParseLanguage, '=', ParseLanguage}

GETPROPERTY = 'GETPROPERTY*', ParseLanguage, '|', ParseLanguage

SETPROPERTY = 'SETPROPERTY*', ParseLanguage, '|', ParseLanguage, '=', ParseLanguage

DRAW = 'DRAW*', ParseLanguage

ANSWER = 'ANSWER*', ParseLanguage

PLAYSOUND = 'PLAYSOUND*', ParseLanguage

DELAY = 'EVENT*', IntegerLiteral

EVENT = 'EVENT*', BooleanExpression

INCORRECTEVENT = 'INCORRECTEVENT*', BooleanExpression

GRASP = 'GRASP*', ParseLanguage

UNGRASP = 'UNGRASP*', ParseLanguage

ExtendedObject = ParseLanguage, '#', IntegerLiteral

Assignment = ParseLanguage, '=', ParseLanguage, {'|', ParseLanguage}

RandomAssignment = ParseLanguage, '~', ParseLanguage, {'|', ParseLanguage}

NumericExpression = IntegerLiteral | ('-', NumericExpression) | (NumericExpression, ('+' | '-' | '*' | '/'), NumericExpression)

IntegerLiteral = ('1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9'), {'0' | '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9'}

BooleanExpression = ('NOT', BooleanExpression) | (BooleanExpression, ('or' | 'and'), BooleanExpression) | ('(', BooleanExpression, ')') | IntersectExpression | SubsetExpression

IntersectExpression = 'INTERSECT(', ParseLanguage, ')'

SubsetExpression = 'SUBSET(', ParseLanguage, ')'

В данной главе также описан синтаксис языка TaskLang, подробно рассмотрена процедура генерации случайных заданий.

В приложении 2 приведено руководство к изучению языка TaskLang.

Таким образом, алгоритмический язык программирования TaskLang, обладающий достаточно простым синтаксисом, создан с целью дать возможность самим специалистам-филологам, не знающим языков программирования, создавать различные задания для обучения и проверки знаний по языкам. В самом языке предусмотрено формирование случайных заданий, что позволяет наилучшим образом понять пользователю смысл

изучаемого понятия, наиболее объективно оценить это понимание на экзамене и облегчить процесс написания однотипных заданий их составителю.

В четвёртой главе приведено описание автоматизированной системы Language Education System (LES), созданной для операционного представления понятий естественных языков и разработки на основе такого представления независимых обучающих и комплексных экзаменуемых приложений по естественным языкам. Данная автоматизированная система состоит из редактора заданий и тестирующей программы (экзамена). Приложения имеют модульную структуру. Каждый модуль является обособленной частью набора типов заданий, которые требуются для того, чтобы полностью написать учебник или экзамен по языку. Редактор и экзамен используют в своей работе общую библиотеку модулей, которая в свою очередь использует два основных модуля – модуль текстовых заданий (генератор текста) и модуль графических заданий (генератор сцен). Модуль текстовых заданий использует реализацию интерпретатора языка TaskLang, а модуль графических заданий, в свою очередь, использует реализацию интерпретатора языка TaskLang с поддержкой графических функций.

С автоматизированной системой LES работают два вида пользователей: составитель экзамена и экзаменуемый человек.

Составитель имеет возможность составить собственный экзамен (рис. 1), то есть: дать название экзамену; задать время, в течение которого сдаётся экзамен; составить задание какого-либо типа; сформировать дерево из составленных заданий. Составитель также проводит экзамен и контролирует результаты путём просмотра отчёта и выставления итоговой экзаменационной оценки.

Экзаменуемый человек может сдать экзамен по языку (рис. 1), выполняя задания, и ознакомиться с результатами экзамена после его завершения.

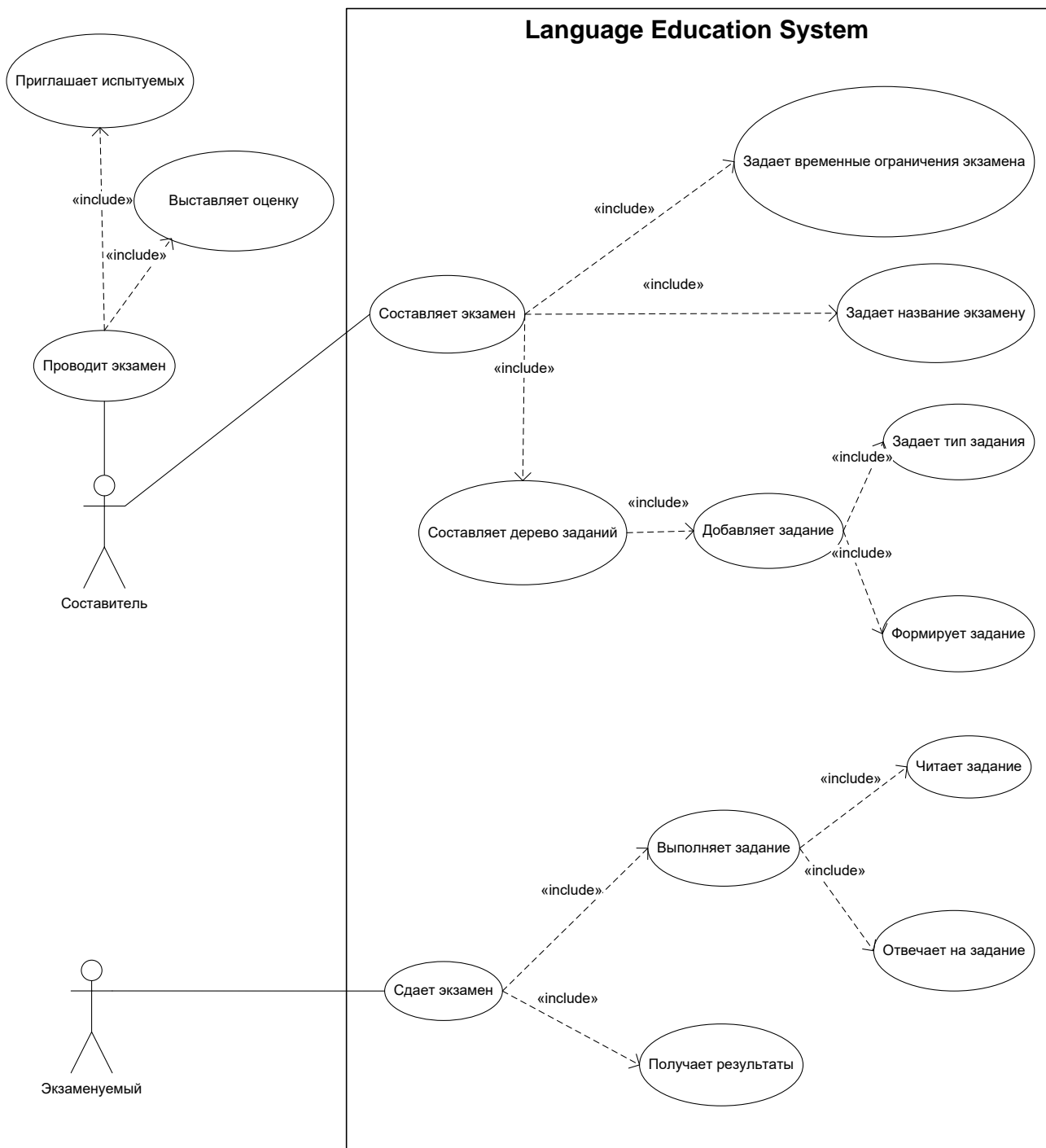


Рис. 1. Работа пользователей с LES

Редактор модулей – это часть автоматизированной системы, предназначенная для написания заданий на языке TaskLang.

В редакторе есть шаблоны типов заданий, с помощью которых составителю легче разрабатывать задания. Разработаны шаблоны для таких типов заданий, как диктант, текст, действия. Каждый шаблон реализован в виде отдельного модуля.

Задание типа диктант имеет своей особенностью то, что в нём обязательно присутствует звуковая составляющая. К этому типу относятся задания, где требуется написать произнесенное программой слово, ответить на

аудиовопрос и др. В задание типа диктант можно включать звуковые файлы, текст, графические файлы.

Задание типа текст в отличие от диктанта не содержит звуковых файлов, задание даётся с помощью текста. К такому типу относятся, например, задания: вставить пропущенные буквы, записать слово в требуемой форме (падеж, число, время и т. д.), понять текст и ответить на поставленные вопросы и др.

Задания типа действие требуют от пользователя выполнения некоторых действий с объектами на сцене с помощью курсора мыши.

Тестирующая программа – это часть автоматизированной системы, являющаяся экзаменом (тестом), составленным в редакторе модулей. В главе 4 приводится подробное описание работы пользователей-экзаменуемых с тестирующей программой LES.

В данной главе приведено также описание основных классов, разработанных в автоматизированной системе LES, и тестирование LES.

В приложении 1 приведена сопроводительная документация к автоматизированной системе LES.

ВЫВОДЫ

В работе получены следующие научные и практические результаты:

1. Предложено определение математической модели понятия естественного языка, на основе которой строится компьютерная модель данного понятия.
2. Определены геометрические параметры пространств, в которых реализуются понятия естественного языка. Установлена алгебраическая и топологическая структура для определенных классов глаголов. Действия данных глаголов сохраняют определённые инварианты пространств, в которых эти глаголы реализуются. На пространстве объектов вводится топология, соответствующая хаусдорфовой метрике.
3. Преобразования моделируемых объектов определены динамической системой дифференциальных уравнений с внешним воздействием пользователя. Перемещение объектов на дисплее осуществляется с помощью движения курсора, которое преобразуется в движение объекта при помощи данной системы дифференциальных уравнений. Получена конечно-разностная схема, аппроксимирующая данную систему дифференциальных уравнений.
4. На основе описанных выше результатов для некоторых понятий кыргызского языка построены их математические модели с учётом пространств, в которых эти понятия реализуются.
5. Разработан язык программирования TaskLang, с помощью которого специалисты-филологи, не владеющие языками программирования, могут создавать учебники и экзамены по естественным языкам. В языке TaskLang предусмотрено формирование случайных заданий.

6. Построена автоматизированная система Language Education System (LES), предназначенная для создания независимых обучающих и комплексных экзаменуемых приложений по естественным языкам.
7. С помощью автоматизированной системы LES созданы первая версия комплексного экзамена и основа независимого электронного учебника по кыргызскому языку.

Автоматизированная система LES одобрена Департаментом по государственному языку Кыргызской Республики. Комплексный экзамен по кыргызскому языку был использован как демонстрационный в ряде вузов (КРСУ, ЖАГУ, МУК, КГУСТА и др.), школ и в Институте теоретической и прикладной математики НАН КР. С помощью автоматизированной системы LES разработан программный продукт для создания учебников «Интерактивная инструментальная среда изучения естественного языка». Автоматизированная система LES апробирована и используется в средних и высших учебных заведениях Бишкека; в приложении 3 даны соответствующие копии актов о внедрении.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

- 1 **Долматова П. С.** Построение дифференциальных уравнений для гибкого управления объектами с помощью компьютерной мыши [Текст] / П. С. Панков, П. С. Долматова // Исследования по интегро-дифференциальным уравнениям, выпуск 37. – Бишкек: Илим, 2007. – С. 18 – 23.
- 2 **Dolmatova P.** Algorithmical Language for Computer-Based Presentation of Notions [Текст] / P. Pankov, P. Dolmatova // ИКЕССО'2007: Proceedings of the 4th International Conference on Electronics and Computer. – Almaty: Suleyman Demirel University, 2007. – Pp. 274 – 279.
- 3 **Долматова П. С.** Математические модели понятий русского языка [Текст] / П. С. Панков, П. С. Долматова // Вопросы преподавания русского языка и литературы в учебных заведениях Кыргызстана. Материалы международной научно-практической конференции. – Москва: Издательство ЦМО МГУ, 2008. – С. 51 – 55.
- 4 **Dolmatova P.** Algorithmic Language and Classification of Verbs for Computer-Based Presentation [Текст] / P. Pankov, P. Dolmatova // Academic Review, № 1 (7). – American University of Central Asia, Bishkek, 2008. – Pp. 233 – 239.
- 5 **Dolmatova P.** Geometrical Problems in Interactive Computer Presentation of Notions of Natural Languages [Текст] / P. Pankov, P. Dolmatova // Actual Problems of Control Theory, Topology and Operator Equations. International Jubilee Conference at the Kyrgyz-Russian Slavic University. – Shaker Verlag, Aachen, Germany, 2009. – Pp. 85 – 87.
- 6 **Dolmatova P.** Software for Complex Examination on Natural Languages [Текст] / P. Pankov, P. Dolmatova // Human Language Technologies as a Challenge for Computer Science and Linguistics: Proceedings of the 4th

- Language and Technology Conference. – Poznan, Poland, 2009. – Pp. 502 – 506.
- 7 **Долматова П. С.** Пространства и преобразования, соответствующие глаголам естественных языков [Текст] / П. С. Долматова // Международный научный журнал «Наука. Образование. Техника», № 2 (29), часть 1. – КУУ, Ош, 2009. – С. 22 – 24.
 - 8 **Долматова П. С.** Разработка программного обеспечения для составления экзаменационных заданий по языкам [Текст] / П. С. Долматова // Материалы итоговой научно-практической конференции «Сотрудничество и связи в сфере высшего образования стран Европы и Центральной Азии в области информационных и инновационных технологий». – Бишкек: КНУ им. Ж. Баласагына, 2009. – С. 141 – 145.
 - 9 **Долматова П. С.** Использование компьютера для самостоятельного изучения и проверки знаний по естественным языкам [Текст] / П. С. Долматова // Известия Кыргызской Академии образования, № 4 (12) – Бишкек, 2010. – С. 283 – 287.
 - 10 **Долматова П. С.** Разработка программного обеспечения для составления обучающих и проверочных заданий по естественным языкам [Текст] / П. С. Долматова // Журнал «Математические структуры и моделирование», выпуск 22. – ОмГУ, Омск, 2011. – С. 74 – 85.
 - 11 **Долматова П. С.** Автоматизированная система для разработки комплексных контролирующих программных средств по естественным языкам [Текст] / П. С. Долматова // Вестник КНУ им. Ж. Баласагына, юбилейный выпуск. – Бишкек, 2012. – С. 458 – 462.
 - 12 **Долматова П. С.** Автоматизированная система для построения независимого интерактивного представления понятий естественных языков [Текст] / П. С. Долматова // Журнал «Проблемы информатики и управления», №1. – Бишкек, 2012. – С. 89 – 98.

КОРУТУНДУ

Долматова Полина Сергеевна

«Табиғый тилдердин түшүнүктөрүн интерактивдүү операциялык көрсөтүүнүн автоматташтырылган системасы» диссертациясы 05.13.06 – технологиялык процесстерди жана өндүрүштөрдү автоматташтыруу жана башкаруу адистиги боюнча техникалык илимдердин кандидаты окумуштуулук даражасын алуу үчүн сунушталган

Урунттуу сөздөр: табиғый тил, комплекстик текшерип көрүүчү программалык колдономо, көз карандысыз үйрөтүүчү программалык колдономо, математикалык модель, кокус тапшырма, алгоритмдик тил, автоматташтырылган система.

Изилдөө объекти: компьютер менен табиғый тилдерди үйрөтүүнү процесси.

Изилдөөнүн максаты: табигый тилдердин түшүнүктөрүн интерактивдүү аракеттөөчү чагылдыруу жана бул чагылдыруунун негизинде табигый тилдер боюнча көз карандысыз үйрөтүүчү жана комплекстик текшерип көрүүчү программалык колдонмолорду түзүү үчүн автоматташтырылган системасын жасоо.

Изилдөөнүн ыкмалары: табигый тилдердин түшүнүктөрүндү көз карандысыз чагылдыруу ыкмасы, объектке-багытталган программалоо ыкмалары.

Изилдөөнүн негизги жыйынтыктары жана алардын жаңылыгы. Табигый тилдин түшүнүгүнүн математикалык моделдинин сунуш кылынган аныктамасынын негизинде бул түшүнүгүнүн компьютердик модели түзүлөт. Түшүнүктөр жүзөгө ашырылганы үчүн мейкиндиктердин геометриялык параметрлери аныкталган. Моделдештирилген объекттерди өзгөртүүлөр болгон дифференциалдык (айырмалык) теңдемелер динамикалык система катары аркылуу аткарылат. Түшүнүктү көз карандысыз компьютердик чагылдыруу үчүн формалдуу эмес алгоритмдик тил курулган; андан кийин бул тил *TaskLang* формалдык алгоритмдик тилинин түрүндө жүзөгө ашырылган. Ал тил табигый тилдер боюнча үйрөтүүчү жана текшерип көрүүчү программалык колдонмону түзүүнүн процессин автоматташтырууга мүмкүнчүлүк берет. Кокус тапшырмалардын, ошондой эле дисплейде объекттер менен аракеттерди пайдалануучудан талап кылган тапшырмалардын түзүүнүн *TaskLang* тилинде мүмкүндүк бар. Автоматташтырылган системанын жардамы менен кыргыз тили боюнча комплекстик электрондук экзамен жана көз карандысыз электрондук окуткучтун негизи түзүлгөн.

Изилдөөнүн жыйынтыктарын колдонуу: автоматташтырылган система Кыргыз Республикасынын орто жана жогорку окуу жайларында киргизилген; тилдер боюнча автордук окуулукарды же экзамендерди түзүү үчүн тилчилер аны пайдаланышат.

Колдонуу аймагы: тилдер боюнча автордук окуткучтарды же экзамендерди түзүү үчүн тилчилер автоматташтырылган системаны колдоно алат.

РЕЗЮМЕ

диссертации Долматовой Полины Сергеевны на тему:

«Автоматизированная система интерактивного операционного представления понятий естественных языков» на соискание ученой степени кандидата технических наук по специальности 05.13.06 – автоматизация и управление технологическими процессами и производствами

Ключевые слова: естественный язык, комплексное экзаменуемое приложение, независимое обучающее приложение, математическая модель, случайное задание, алгоритмический язык, автоматизированная система.

Объект исследования: процесс обучения естественным языкам с помощью компьютера.

Цель исследования: разработка автоматизированной системы интерактивного операционного представления понятий естественных языков и создания независимых обучающих и комплексных экзаменуемых приложений по естественным языкам.

Методы исследования: метод независимого представления понятий естественных языков, методы объектно-ориентированного программирования.

Основные результаты работы и их новизна. Предложено определение математической модели понятия естественного языка, на основе которой строится компьютерная модель данного понятия. Определены геометрические параметры пространств, в которых реализуются понятия естественного языка. Преобразования моделируемых объектов производятся динамической системой дифференциальных (разностных) уравнений. Построен неформальный алгоритмический язык для независимого компьютерного представления понятий, далее реализованный в виде формального алгоритмического языка *TaskLang*. Этот язык позволяет автоматизировать процесс разработки обучающих и экзаменуемых приложений по естественным языкам. В язык *TaskLang* заложена возможность создания случайных заданий, включая задания, требующие от пользователя действий с объектами на дисплее. С помощью автоматизированной системы создан комплексный экзамен и основа независимого электронного учебника по кыргызскому языку.

Использование результатов исследования: автоматизированная система внедрена в средних и высших учебных заведениях Кыргызской Республики и используется специалистами-филологами для создания авторских учебников или экзаменов по языкам.

Область применения: автоматизированная система может быть использована специалистами-филологами для составления авторских учебников и экзаменов по языкам.

RESUME

Dolmatova Polina Sergeyevna «Computer-based System for Interactive Operational Representation of Natural Languages' Notions» for the scientific degree of Ph.D. (Candidate of Technical Sciences) in the specialty 05.13.06 – Automation and Control of Technological Processes and Production

Keywords: natural language, integrated testing application, independent teaching application, mathematical model, random assignment, algorithmic language, computer-based system.

Object of the research: process of studying natural languages with a computer.

Objective of the research: to develop a computer-based system for interactive operational representation of natural languages' notions and development

independent teaching and integrated testing applications on natural languages based on the above representation.

Methods of the research: method of independent representation of natural languages' notions, methods of the object-oriented programming.

Basic results of the work and their novelty. The definition of the mathematical models of natural language's notion was suggested; the computer model of this notion is built on its mathematical model. The geometric parameters of the spaces were defined which natural language's notions are implemented in. Transformations of simulated objects were defined by the dynamic system of differential (difference) equations. The informal algorithmic language for independent computer representation of notions based on the definition of computer model of notion is proposed; then implemented as the formal algorithmic language *TaskLang*. *TaskLang* enables to automate development of teaching and testing applications on natural languages. *TaskLang* contains the possibility for creating random assignments, including assignments that require user actions with the objects on display. With this computer-based system the integrated examination and the basis of the independent electronic textbook on Kyrgyz were created.

Use of the results of the research: the computer-based system is implemented in secondary and higher education institutions of the Kyrgyz Republic and used by philologists for creating author's textbooks or exams on languages.

Scope: the computer-based system can be used by philologists for creating author's textbooks or exams on languages.

Долматова Полина Сергеевна

**АВТОМАТИЗИРОВАННАЯ СИСТЕМА ИНТЕРАКТИВНОГО
ОПЕРАЦИОННОГО ПРЕДСТАВЛЕНИЯ ПОНЯТИЙ ЕСТЕСТВЕННЫХ
ЯЗЫКОВ**

Автореферат диссертации

Подписано к печати 14.12.2012г. Формат бумаги 60x84¹/₁₆.

Бумага офс. Печать офс. Объем 1,5 п.л. Тираж 120 экз.

г.Бишкек, ул, Сухомлинова, 20. ИЦ “Текник” КГТУ, т.: 54-29-43

E-mail: beknur@mail.ru

