

**УЗБЕКСКОЕ АГЕНСТВО СВЯЗИ И ИНФОРМАТИЗАЦИИ
ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

На правах рукописи

АХАТОВА РАШИДА ЮНУСОВНА

**ТЕХНОЛОГИЯ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ
ТЕСТИРОВАНИЯ И ЕЕ ВНЕДРЕНИЕ В УЧЕБНЫЙ ПРОЦЕСС**

Специальность: 5A140901 – «Профессиональное образование»

(информатика и информационные технологии)

ДИССЕРТАЦИЯ

на соискание академической степени магистра

Работа рассмотрена
и допускается к защите
зав. кафедрой «Педагогика технического
образования»

к.т.н. Мигранова Э. А. _____

«___» _____ 2012 г.

Научный руководитель:

д.ф-м.н., академик Абуталиев Ф. Б. _____

«___» _____ 2012 г.

Ташкент 2012

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

ГЛАВА 1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ

1.1. Интеллектуальные обучающие системы

1.1.1. Архитектура интеллектуальной обучающей системы

1.1.2. Обзор существующих интеллектуальных обучающих систем

1.2. Исследование технологии построения интеллектуальной системы тестирования

1.2.1. Сравнительный анализ характеристик существующих интеллектуальных систем тестирования

1.2.2. Структура интеллектуальной системы тестирования

1.3. Постановка задачи

Выводы к 1 главе

ГЛАВА 2. ТЕХНОЛОГИЯ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ

2.1. Модели построения интеллектуальной системы тестирования

2.1.1. Модель распознавания образа уровня знаний

2.1.2. Модель предметно - критериального составления тестов

2.1.3. Модель Раша

2.2. Технология интеллектуального тестирования

2.3. Разработка концепции интеллектуальной системы тестирования на примере модели Раша

Выводы ко 2 главе

Глава 3. АПРОБАЦИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ В УЧЕБНОМ ПРОЦЕССЕ

3.1. Проведение педагогического эксперимента

3.2. Анализ результатов и эффективности интеллектуальной системы тестирования

Выводы к 3 главе

Заключение

Список использованной литературы

Приложения

ВВЕДЕНИЕ

Развитие и широкое применение информационных технологий является глобальной тенденцией мирового развития и научно-технической революцией последних десятилетий.

Современное общество характеризуется интенсивным участием в процессе информатизации. Информатизация общества - это процесс овладения информацией как ресурсом управления и развития на базе средств информатики, включающей медиатизацию, компьютеризацию и интеллектуализацию.

Сегодня информатизация как технико-технологическая база становления информационного общества выступает национальным стратегическим ресурсом, определяющим не только общий уровень социального и культурного развития государства, но и его место в глобальном процессе мирового развития.

Согласно Указа Президента Республики Узбекистан И. А. Каримова "О дальнейшем развитии компьютеризации и внедрении информационно-коммуникационных технологий", и в соответствии с Национальной программой по подготовке кадров Республики Узбекистан, необходимо создать национальную систему информатизации и ее массовое внедрение во все сферы жизни общества, создание благоприятных условий для вхождения в мировое информационное сообщество [1].

Основной задачей системы образования является удовлетворение потребностей государства в выполнении социального заказа. Современные запросы влекут за собой непрерывное совершенствование учебного процесса, поэтому процесс обучения должен быть достаточно гибким для быстрой адаптации к меняющимся требованиям. Одним из средств, способствующих повышению качества подготовки учащихся и формированию соответствующей образовательной среды, являются новые информационные технологии обучения, базирующиеся на применении обучающих систем с элементами искусственного интеллекта. При разумной организации они интенсифицируют процесс обучения, обеспечивают формирование глубоких знаний, выработку прочных умений и твердых навыков, а также вносят свой вклад в процесс воспитания учащихся.

В настоящее время процесс обучения нуждается в серьезном совершенствовании и научном обосновании принимаемых решений. Это особенно актуально в условиях все возрастающих требований к подготовке обучающихся, необходимостью обновления учебных планов и учебно-методического обеспечения, необходимости повышения качества учебного процесса в условиях современной жизни. Необходим поиск новых

подходов, обеспечивающий целесообразную перестройку системы образования с учетом жизненных реалий.

Наиболее перспективными средствами разрешения этих проблем являются инструментальные системы, которые принято называть обучающими системами с элементами искусственного интеллекта.

Актуальность работы. Интеллектуальная система тестирования позволяет получать объективные оценки знаний испытуемых по единой нормированной шкале, сопровождаемые качественными характеристиками. В этой связи разработка единых алгоритмов на интеллектуальной базе позволяет эффективным и экономичным путем достигать высоких результатов тестирования.

Объектом исследования является процесс интеллектуального тестирования на примере курса «Мультимедийные технологии в образовании».

Предметом исследования является исследование технологии разработки интеллектуальной системы тестирования.

Целью работы является исследование технологии разработки интеллектуальной системы тестирования. Данная цель предполагает решение следующих **задач**:

- на основе анализа литературных источников рассмотреть архитектуру и модели построения существующих интеллектуальных систем тестирования;
- разработка технологии интеллектуальной системы тестирования, с использованием модели Раша;
- апробация технологии интеллектуальной системы тестирования;
- проведение экспериментальных исследований и оценка их эффективности.

Методы исследования основаны на использовании методов системного анализа, планирования экспериментов, построения моделей обучаемого, теории вероятностей и математической статистики, оптимизации параметрического тестирования, анализа психологических показателей обучаемого.

Научная новизна работы состоит в следующем:

1. Предложена технологическая модель интеллектуальной системы тестирования, направленная на индивидуализацию подхода к каждому обучаемому, путем целевого выделения наиболее ориентированных частей системы.

2. Разработана технологическая структура интеллектуальной системы тестирования и ее возможные архитектурные преобразования в зависимости от поставленных целей.

3. Предложены методические рекомендации по применению интеллектуальной системы тестирования для оценки знаний обучаемых.

Практическая значимость результатов исследования заключается в разработке соответствующего алгоритмического и методического обеспечения, доступного для реализации в различных интеллектуальных системах тестирования.

Апробация работы. Основные теоретические и практические результаты диссертационной работы докладывались и обсуждались на Республиканской научно-методической конференции “Современные информационные технологии в телекоммуникации и связи”, посвященная 100-летию со дня рождения Исламова Анвара Исламовича (Ташкент, 2011г.); на Республиканской научно-методической конференции одаренных студентов, магистрантов и аспирантов “Инновационные технологии и проблемы их теоретического и практического применения в учебном процессе” (Ташкент, 2011г.).

Структура и объем диссертации. Работа состоит из введения, трех глав с выводами, заключения, списка использованной литературы из 40 наименований и приложения.

Краткое содержание диссертации.

В первой главе рассмотрены классификация и особенности интеллектуальных обучающих систем, приводится анализ существующих подходов к их построению; исследовано современное состояние процесса создания тестирующих систем и проблемы построения интеллектуальной системы тестирования.

Во второй главе были проанализированы существующие методы и модели, позволяющие построить интеллектуальную систему тестирования. Разработана технология интеллектуальной системы тестирования на примере модели Раша.

В третьей главе рассмотрена апробация технология интеллектуальной системы тестирования в учебном процессе на примере курса «Мультимедийные технологии в образовании», приведены результаты экспериментов, на основании которых доказана эффективность применения предлагаемой технологии.

ГЛАВА 1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ

1.1 Интеллектуальные обучающие системы

Известно, что системы, использующие методику искусственного интеллекта, называют интеллектуальными обучающими системами (ИОС). Интеллектуальные обучающие системы - это качественно новая технология, отличительными особенностями которой являются: моделирование процесса обучения с использованием динамически развивающейся базы знаний; автоматический подбор рациональной стратегии обучения для каждого обучаемого; автоматизированный учет новой информации, поступающая в базу данных. ИОС реализует адаптивное и двухстороннее взаимодействие, направленное на эффективную передачу знаний. Здесь, под адаптивностью понимается то, что система дает персональные пояснения, для каждого обучаемого, с помощью динамического управления, зависящего от процесса обучения. Двустороннее взаимодействие - это взаимодействие со смешанной инициативой, при которой обучаемый может задать вопросы или с помощью системы решить задачу. В основном, существующие ИОС отличаются друг от друга, прежде всего методологиями представлений знаний о предметной области, об обучаемом и о процессе обучения [3].

Анализ показывает, что наиболее перспективным путем развития ИОС является создание самообучающихся систем, приобретающих знания в диалоге с человеком.

Основным недостатком многих существующих ИОС является ориентация на специальные знания в рамках определенного предмета, так что в них не предусмотрена возможность простой адаптации к другой предметной области. Далее, более общий подход будет состоять в развитии интеллектуальной оболочки, из которой затем можно получить различные ИОС путем наполнения соответствующим содержанием, как базы знаний.

В современных интеллектуальных обучающих системах, в основном, используются знания о качественных и количественных аспектах процесса обучения. Однако необходимо учитывать и мотивационную сторону обучения. Мотивационные аспекты обучения можно классифицировать в соответствии с такими явлениями, как соревновательность, заинтересованность, самоконтроль, уверенность и удовлетворение.

Поэтому обучающая система должна:

- определять мотивационное состояние обучаемого;
- реагировать с целью мотивации рассеянных, менее уверенных или недовольных учеников, или поддержки уже мотивированных учеников. Примеры мотивационной тактики:
 - если менее уверенный ученик правильно решает задачу, система может предложить ему подобную задачу для закрепления;
 - внимание рассеянных или неактивных обучаемых может быть привлечено неожиданными эффектами или вводными комментариями;
 - интерес может быть повышен головоломками, вопросами или знакомством с новыми темами.

1.1.1. Архитектура интеллектуальной обучающей системы

Обучающие программы служат для обеспечения следующих педагогических целей:

1) **демонстрация учебного материала:** обучаемому предъявляется текстовой, графической, аудио- и видеоматериал в некоторой фиксированной последовательности. Программы, реализующие эту цель, называются демонстрационными, или компьютеризированными лекциями и электронными книгами;

2) **тестирование и диагностика уровня знаний:** обучаемый подвергается испытанию с целью выяснения некоторых его характеристик, в частности, глубины усвоения знаний, степени приобретения навыков и умений, полученных им ранее;

3) **тренировка:** обучаемый получает знания, навыки и умения, необходимые для выполнения определенной работы, а учебные программы организуют адекватную среду для приобретения и закрепления требуемых навыков и умений. Учебные программы данного класса принято называть тренирующими программами или тренажерами. Характерными их чертами являются: а) наличие как формальной, так и аудиовизуальной модели изучаемого предмета; б) явное знание целей функционирования системы; в) наличие контроля действий обучаемого; г) наличие жестких ограничений на время, выделяемое обучаемому для достижения поставленной цели;

4) **обучение:** обучаемый получает знания, навыки и умения в некоторой предметной области под управлением учебной программы, при этом программа принимает на себя все функции преподавателя по организации предъявления учебного материала, контроля его усвоения и диагностики ошибок обучаемого (Рис.1). Данный класс программ характеризуется: а) наличием цели обучения; б) реализацией некоторого метода обучения,

приводящего к достижению требуемой цели и определяющего стиль общения с обучаемым; в) комплексным решением задач обучения, контроля и диагностики уровня знаний [4].

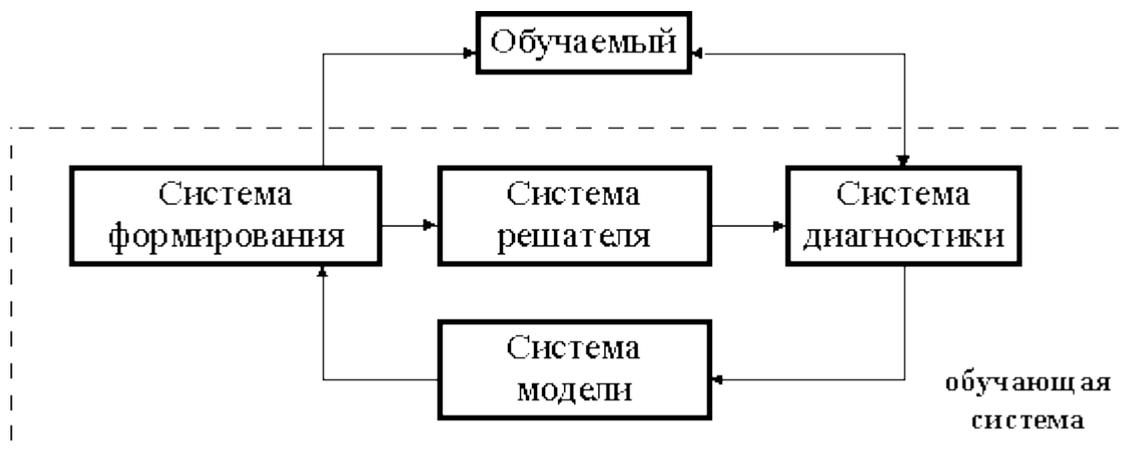


Рис.1. Структура интеллектуальной обучающей системы

Обучающие программы, разработанные с помощью традиционных средств компьютерной технологии обучения, имеют либо демонстрационный тип, чем обеспечивают относительную свободу обучаемого, либо жесткую структуру управления. В отличие от них интеллектуальные обучающие системы дают обучаемому большую свободу, что позволяет обучаемому выбирать учебные темы и задачи, задавать интересующие его вопросы и получать на них ответы, а также получать адекватную диагностику своих знаний и умений. Это внешнее различие основано на том, что в интеллектуальных обучающих системах явно представлены значения об изучаемом предмете и метазнания о процессе управления обучением и состоянии знаний обучаемого.

Основные задачи, возникающие при создании ИОС:

- представление и обработка знаний в предметной области, составляющее предмет обучения;
- разработка методов контроля и диагностики ошибок обучаемого;
- разработка методов управления обучением.

В основе архитектуры ИОС лежит следующая модель процесса обучения: имеется цель обучения, выраженная в терминах текущих характеристик обучаемого, и пока цель не будет достигнута, повторяется следующая последовательность действий:

- на основании текущего состояния обучаемого и методики обучения генерируется очередная задача (здесь задача понимается широко как любая информация, требующая ответных действий обучаемого);
- ответ обучаемого сравнивается с эталонным решением и на основании сравнений различий производится диагностика ошибок обучаемого;

- по результатам диагностики корректируются текущие характеристики обучаемого.

В соответствии с данной моделью процесса обучения ИОС можно рассматривать как совокупность трех взаимодействующих экспертных систем (ЭС):

- ЭС по решению задач в изучаемой предметной области (ПО);
- ЭС по диагностике ошибок обучаемого;
- ЭС по планированию процесса управления обучением.

В структуре ИОС можно выделить следующие базы знаний (БЗ): учебная БЗ для данной ПО; модель обучаемого (МО); БЗ о возможных ошибках обучаемого; БЗ о процессе обучения.

Учебная база знаний (УБЗ) описывает не только основные понятия и методы решения задач в ПО, но и содержит определения понятий, описания методов, примеры, упражнения и задачи. В отличие от базы знаний ЭС по решению задач УБЗ должна явно отражать структуру ПО и стратегические знания о методах решения задач. С другой стороны, УБЗ можно рассматривать как представление ограничений, в рамках которых ЭС планирует обучение [7].

МО содержит информацию о состоянии знаний обучаемого, как общие, интегрированные характеристики, так и те, которые отражают усвоение им текущего учебного материала. Первоначально МО формируется во время предварительного тестирования обучаемого. В терминах МО выражается цель обучения.

База знаний об ошибках обучаемого (БЗО) содержит каталог возможных ошибок обучаемого и правила выдвижения и проверки гипотез, о неправильных представлениях обучаемого, приведших к данной ошибке, на основе различий между решениями, предложенными обучаемым и ЭС, а также текущего состояния МО.

Таблица 1

Типы интеллектуальных обучающих систем

Тип	Состав	Назначение
Консультационная	Учебная среда Объяснение	Консультации при решении задач и поиске учебной информации
Диагностирующая	Решатель задач Диагностика Модель обучаемого	Диагностика ошибок при решении задач
Управляющая	Решатель задач Диагностика	Обучение понятиям и умениям (навыкам) на

	Управление обучением Модель обучаемого	основе моделирования знаний обучаемого
Сопровождающая	Инструментальная система Диагностика Управление обучением Модель пользователя	Слежение за поведением пользователя и помощь при ошибочных или нерациональных действиях

База знаний о процессе обучения содержит знания о планировании и организации процесса обучения, общих и частных методиках обучения.

Для поддержания "разумного" процесса обучения педагог использует специальные знания трех основных типов: о предмете обучения, о стратегии и методах обучения, об обучаемом. К этим специальным типам знаний можно добавить коммуникативные функции: передача знаний обучаемому, выбор формы и других характеристик сообщений, восприятие его ответов. В традиционных автоматизированных системах обучения (АОС) фрагменты знаний, необходимые для реализации конкретной части курса обучения, были жестко встроены в текст отдельных элементарных шагов этого курса. В интеллектуальных обучающих системах необходимые знания явно выделены и представлены, как правило, с помощью методов и технологий инженерии знаний. Используя эти знания, ИОС выполняет различные функции педагога (помощь в процессе решения задач, определение причин ошибок обучаемого, выбор оптимального учебного воздействия). Идеальная ИОС должна "уметь" представлять и использовать все перечисленные типы знаний. Это позволяет определить ее структуру в виде набора взаимодействующих модулей, в каждом из которых сосредоточены знания одного типа [10].

Модуль-эксперт предметной области содержит фактографические, процедурные и другие знания о предметной области (ПрО).

Компонента ИОС, представляющая знания системы об обучаемом, называется моделью обучаемого. Эта модель постоянно обновляется в ходе обучения в соответствии с изменением отражаемых ею характеристик обучаемого. В идеальной ИОС такая модель используется всеми модулями системы для адаптации их работы к конкретному обучаемому. Функции поддержки модели в актуальном состоянии вместе с самой МО могут быть объединены под названием "модуль обучаемого".

Педагогический модуль (модуль-педагог) представляет знания эксперта-педагога об

организации и поддержке целенаправленного процесса обучения. На основе стратегии обучения с учетом модели обучаемого модуль-педагог обеспечивает управление учебной деятельностью. Модуль - интерпретатор содержит в себе знания, необходимые для поддержки взаимодействия с обучаемым.

Модуль обучаемого — центральный компонент, позволяющий обеспечивать качество управления процессом учения. Он реализуется как структура данных, отражающая в каждый момент времени существенные для системы характеристики конкретного обучаемого. При этом МО распадается на две основные компоненты:

- модель личностных характеристик,
- модель знаний о предмете.

Первая отражает достаточно устойчивые личностные (когнитивные) характеристики обучаемого, вторая - картину знаний и умений конкретного обучаемого по изучаемому курсу в данный момент времени.

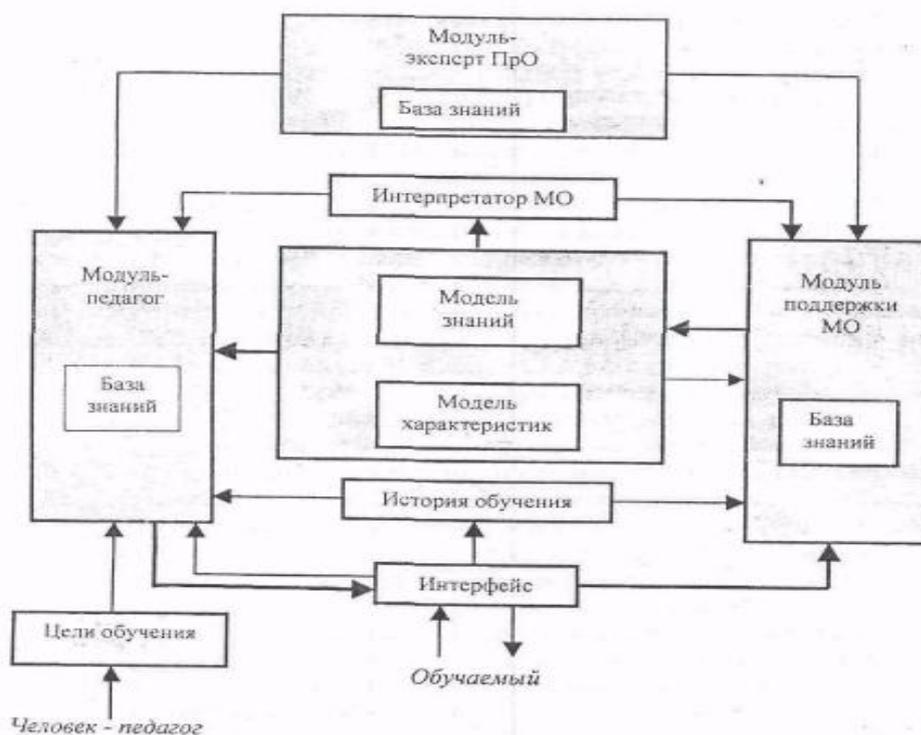


Рис.2. Обобщенная архитектура ИОС

Модель обучаемого после создания ИОС должна поддерживаться в актуальном состоянии. Основной функцией модуля поддержки МО является определение текущего состояния знаний обучаемого на основе его наблюдаемого поведения. В процессе логического вывода для принятия решений по управлению обучением модуль поддержки может использовать не только последние действия обучаемого, но и полную запись вза-

имодействия обучаемого с системой, накапливаемую за всю историю обучения. Решение об уровне текущего состояния знаний конкретного обучаемого модуль поддержки использует собственную базу знаний об обучаемых. В литературе её называют "родовая" модель обучаемого, назначение которой накапливать и поддерживать такие сведения об обучаемом, как типовые ошибки, спектр личностных характеристик и т.д.

Кроме того, модуль поддержки может использовать возможности модуля-эксперта ПроО для сравнения поведения обучаемого с поведением эксперта в той же ситуации, что обеспечивает эффективную "диагностику" знаний.

Текущее состояние знаний и личностные характеристики обучаемого, используются модулем-педагогом для индивидуализации процесса обучения. В некоторых системах МО снабжена машиной логического вывода, которая может моделировать (имитировать) поведение обучаемого. Это позволяет модулю-педагогу предсказывать действия обучаемого, а модулю поддержки МО подбирать соответствующую модель обучаемого.

Используя МО и цели обучения, модуль-педагог выбирает и реализует наиболее оптимальное учебное воздействие. Поддерживается история обучения для обновления состояния знаний обучаемого. Новое состояние МО снова используется модулем-педагогом и т. д. [15].

1.1.2. Обзор существующих интеллектуальных обучающих систем

На сегодняшний день создано достаточное количество интеллектуальных обучающих систем различной классификации и отличающихся функциональными возможностями.

Обучающие системы АОС-ВУЗ, АДОНИС работают по следующей схеме. Обучаемому предлагается порция учебной информации, после чего задаются один или несколько вопросов для проверки качества усвоения материала. В случае неправильного ответа обучаемому выдается либо дополнительная информация и снова задается вопрос, либо подсказка (правильный ответ на вопрос), а затем следующая порция учебной информации. Если последует правильный ответ— сразу переход к следующему шагу обучения. Эти системы ограничиваются оценкой величин, характеризующих уровень знаний обучаемого очередной порции учебной информации, которая используется для определения следующего обучающего воздействия. В дальнейшем эти величины никак не воздействуют на ход обучения.

В рамках технологии создания адаптивных обучающих систем ПАЛЕВАС/И разработан и реализован ряд интеллектуальных обучающих систем: ПАЛЕВАС, АДТ, ИАСОС.

ИОС ПАЛЕВАС ориентирована на обучение пользователей-новичков основам логического программирования, а именно, формализации предметных областей, для которых решается некоторая задача .

Система АДТ ориентирована на специалистов высокой квалификации в области искусственного интеллекта и содержит фрагмент курса по автоматическому доказательству теорем с резолюцией, индукцией и симметрией.

Указанные системы обеспечивают индивидуальное обучение различных категорий пользователей на основе имеющихся знаний об обучаемом, предмете изучения и методике обучения.

Используется модель обучаемого, состоящая из четырех уровней:

- локальный, содержащий информацию о выполнении обучаемым последнего учебного задания;
- текущий, содержащий результаты изучения за текущий период обучения;
- глобальный, содержащий результаты изучения всего курса обучения и последовательность прохождения узлов обучающей сети, отражающей предметную область обучения;
- тестирование, отражающее результаты априорного тестирования.

Локальный уровень используется для выбора очередного обучающего воздействия в рамках одного урока; текущий и глобальный для выбора очередного урока или темы для изучения. Основой для формирования этих уровней модели является анализ работы обучаемого.

В результате априорного тестирования обучающая система получает знания об индивидуальных особенностях обучаемого (возраст, образование, скорость реакции на учебные задания, компетентность при решении задач, тип эмоциональной возбудимости, эмоциональную настроенность на обучение) до того, как обучаемый приступил к занятиям.

В **системе ПАЛЕВАС** обучаемый изучает материал данного урока, затем получает задание первого уровня сложности. В зависимости от выполнения задания обучаемый переходит к выполнению заданий второго уровня сложности либо к выполнению аналогичного задания текущего уровня сложности. Обучаемому на первом уровне дается возможность выполнить три аналогичных задания. Если он с ними не справляется, то переходит на изучение материала по альтернативной ветке (более подробный и легко усваиваемый учебный материал). Если обучаемый хорошо выполнил задания первого и второго уровней сложности, он приступает к выполнению контрольных заданий, отличающихся от учебных тем, что для их выполнения нельзя воспользоваться помощью со стороны системы. Кроме этого, обучаемому перед изучением материала некоторых уроков

предлагается выполнить априорные контрольные задания. Это возникает в том случае, когда для изучения материала текущего урока необходимо знать материал не только предыдущего, но и ранее изученных уроков [18].

В системе АДТ обучаемый переходит к выполнению следующего учебного задания только в том случае, если выполнил предыдущее (количество попыток не ограничено, пока обучаемый не даст правильный ответ), в противном случае обучаемый изучает материал снова с использованием той же учебной информации.

Рассмотрим характерные особенности наиболее известных тьюторских систем, положивших начало проектирования этого класса интеллектуальных систем обучения.

Одной из первых, ныне классических, интеллектуальных тьюторских систем является SCHOLAR, предложенная Карбонеллом в 1970 году и предназначенная для обучения географии Южной Америки студентами университетов.

Интерфейс "обучающая система - обучаемый" в SCHOLAR обеспечивается диалоговым процессом, в котором роли партнеров не зафиксированы. Это означает, что во время обучения вопросы может задавать не только система, но и обучаемый. Для общения с обучаемым SCHOLAR использует ограниченный английский язык. Ограниченность естественного языка, используемого интерфейсом SCHOLAR, заключается в том, что генерируемые вопросы, и фразы не являются произвольными, а создаются при помощи фраз-шаблонов, дополняемых информацией из модели предметной области. Таким образом, отмеченная процедура генерации не предполагает "понимания" естественного языка.

Для ответов на вопросы обучаемого SCHOLAR использует механизм логических умозаключений. Техника логических умозаключений, в большинстве случаев сводится к навигации по узлам сетевой модели предметной области. Например, система может легко прийти к заключению, что город Сантьяго находится в Южной Америке, поскольку узел "Сантьяго" связан с узлом "Чили" отношением "часть-целое", который, в свою очередь, связан тем же типом отношения с узлом "Южная Америка".

В SCHOLAR впервые для моделирования предметной области использовались семантические сети (semantic networks).

Авторы SCHOLAR возлагали большие надежды на семантические сети и ввели понятие модель идеальных знаний обучаемого или модель идеального студента (perfect student) как синоним понятия модели предметной. Под моделью идеального студента в SCHOLAR понималась модель предметной области в виде полной семантической сети. Первоначальная версия SCHOLAR характеризовалась чрезвычайно простой стратегией обучения, которая заключалась в случайном выборе тем. В последующих версиях

развитие стратегии обучения осуществлялось в направлении замены процедуры случайного выбора процедурой, действующей на основе набора эвристических правил, полученных из наблюдений за работой опытных тьюторов. Тема, выбранная либо на основе отмеченных эвристических правил, либо по запросу обучаемого, динамически включалась в план занятия, и ей выделялось время на обсуждение, пропорциональное ее важности.

1.2. Исследование технологии построения интеллектуальной системы тестирования

В процессе освоения некоторой учебной дисциплины обучаемый приобретает или совершенствует знания, умения и навыки. Каждому существенному с точки зрения учебного процесса знанию, умению и навыку можно сопоставить некоторый компонент оверлейной модели, характеризующий степень владения обучаемым данным элементом знаний. Значение каждого компонента модели, в общем случае, является оценкой некоторой нечеткой величины посредством степени принадлежности нечеткому множеству, фактора уверенности или субъективной вероятности. Таким образом, знания обучаемого учебной дисциплины представляются множеством или, в общем случае, сетью покомпонентных оценок. При проведении контроля можно поставить вопросы об: 1) интегральной оценке уровня знаний обучаемого (тестирование или контроль подготовленности), 2) о выявлении тех подобластей учебной дисциплины, знания и/или умения о которых неудовлетворительны (диагностика) [23].

Тестирование и диагностика знаний производится посредством опроса обучаемого. Каждый вопрос связан с одним или несколькими элементами знания об изучаемой дисциплине и обладает определенным весом («важностью», сложностью). Если выбор следующего вопроса зависит от ответа обучаемого на предыдущий вопрос, то тестирование (диагностика) знаний называется адаптивным или условным, а в противном случае неадаптивным или безусловным. Различают также тестирование (диагностику) с безусловной и условной остановкой. В первом случае процесс заканчивается после вывода всех запланированных вопросов, во втором — при достижении обучаемым и/или системой одного из некоторого множества заключительных состояний. Преимущество адаптивного тестирования (диагностики) с условной остановкой заключается, с одной стороны, в том, что задается меньше вопросов и, следовательно, сокращается время тестирования (диагностики), а с другой стороны, вопросы задаются целенаправленно,

следовательно, тестирующая (диагностирующая) система выглядит более разумной и вызывает больше доверия у обучаемого».

Прежде чем дать формальные описания задач тестирования и диагностики знаний, рассмотрим их области применения.

Тестирование используют для контроля:

- 1) предполагаемых предварительных знаний и умений, необходимых для успешной работы с обучающей системой (предварительное тестирование);
- 2) знаний и умений, полученных в результате обучения с помощью ЭВМ, учителя, самостоятельно и пр. (заключительное или итоговое тестирование).

Диагностика используется для:

- 1) определения психологических характеристик обучаемого;
- 2) формирования начального состояния модели обучаемого;
- 3) выработке гипотез о текущем состоянии знаний обучаемого с целью управления процессом учения.

Рассмотрим граф описания предметной области, состоящий из вершин-понятий (умений), соединенных дугами-отношениями типа «быть базовым». Зададим произвольное разбиение множества вершин графа на подмножества, причем допускаются пересечение и строгое включение одних подмножеств в другие. С каждым подмножеством свяжем одну или несколько гипотез о состоянии знаний и умений обучаемого. Например, «обучаемый не знает понятий С1, С2 и связей между ними», «обучаемый не умеет применять теорему Пифагора», «обучаемый владеет предметом на оценку «хорошо»». Как правило, с каждой гипотезой связывают некоторое характеризующее ее значение, которое интерпретируют в случае диагностики как вероятность незнания обучаемым данного понятия (умения), а в случае тестирования — как вероятность владения предметом на оценку, задаваемую гипотезой. В качестве примера рассмотрим два предельных случая разбиения графа описания предметной области. Пусть разбиение состоит из одного множества, включающего весь граф. Сопоставим ему две гипотезы: 1) обучаемый владеет предметной областью (зачет) и 2) обучаемый не владеет ею (незачет). Тогда вероятности гипотез представляют собой статистические данные о владении данным предметом некоторым контингентом обучаемых и могут использоваться как исходные - для определения уровня знаний конкретного обучаемого. Другое предельное разбиение графа получается, когда классы и понятия находятся во взаимно однозначном соответствии. Тогда сопоставив каждому классу гипотезу «обучаемый владеет данным понятием», характеризуемой некоторой

вероятностью, получим оверлейную модель обучаемого [14].

Каждый гипотезе соответствует множество симптомов. Под симптомом понимается некоторое событие, наличие или отсутствие которого влияет на подтверждение гипотезы. В каждую гипотезу симптом входит с некоторым весом, характеризующим его вклад в гипотезу. Симптомы вырабатываются в результате анализа решения задач, причем решение одной задачи может означать несколько симптомов.

Таким образом, задача тестирования или диагностики знаний может быть представлена в виде тройки (H, S, P), где

$H = \{H_i\}$ ($i=1, n$) — множество гипотез,

$S = \{s_j\}$ ($j=1, m$) — множество симптомов,

$P = \{p_k\}$ ($k=1, l$) — множество задач.

Таблица 2

Матрица соответствия гипотеза-симптомы

Симптомы	Гипотезы			
	H_1	H_2	H_n
s_1	w_{11}	w_{21}	w_{n1}
s_2	w_{12}	w_{22}	w_{n2}
....

Таблица 3

Матрица соответствия задача-симптомы

Задачи	Симптомы				Затраты
	s_1	s_2	...	s_m	
p_1	r_{11}	r_{12}		r_{1m}	c_1
p_2	r_{21}	r_{22}		r_{2m}	c_2
p_l	r_{l1}	r_{l2}		r_{lm}	c_l

Каждая гипотеза H_i из H состоит из оценки v_i и вектора m - $\{w_{ij}\}$, где w_{ij} -вес j -й симптомы в i -й гипотезе. Вес симптома в самом простом случае может принимать значение 1, если симптом влияет на гипотезу, и 0 в противном случае. В общем случае вес симптома может быть числом в интервале от 0 до $+M$ (где M , например, равно 100), причем абсолютное значение веса определяет важность вклада симптома в гипотезу, а его знак указывает, увеличивается или уменьшается оценка гипотезы при данном симптоме.

С выполнением каждой задачи pk из P связана оценка затрат ck , например время на подготовку задачи системой и среднее время ее решения. Кроме того с ней связан характеристический вектор $rk = \{rkj\}$, задающий соответствие между задачей и симптомами — rkj , который равен 1, если задача pk означает симптом sj имеет решение, в противном случае, а также некоторый механизм для анализа решения задачи которое представляет обучаемый, и означивания симптомов. В табл. 2 и табл.3 приведены исходные данные задачи тестирования и диагностики знаний. В частном случае, когда имеется взаимно однозначное соответствие между симптомами и задачами, т. е. матрица R является единичной, исходные данные можно описать одной таблицей, получаемой из табл. 2 заменой симптомов соответствующие им задачи и добавив справа вектор-столбец затрат. Такая постановка задачи более традиционна и широко применяется в теории вопросника.

Для тестирования и диагностики знаний применяются два основных подхода, первый состоит в применении оптимизационных методов теории вопросников для синтеза минимальных деревьев принятия решений, а второй — в использовании механизмов принятия решений, применяемых в ЭС. Рассмотрим оба подхода.

Вопросником Q называется дерево поиска (принятия решения), все висячие вершины которого соответствуют гипотезам из H , а каждой внутренней вершине и корню соответствует задача (вопрос) из P . Ценой идентификации гипотезе Hi называют сумму цен вопросов, принадлежащих пути $\Gamma(Hi)$, соединяющего корень дерева с гипотезой Hi :

Цена вопросника определяется как математическое ожидание цены идентификации гипотез

Основной проблемой теории вопросников является разработка алгоритмов построения оптимальных вопросников, причем в качестве критерия оптимальности, как правило, принимается минимизация цены вопросника. Теория вопросников нашла применение для диагностики технических устройств, тестирования программ и тестирования диагностики знаний. Для построения оптимальных вопросников применяются методы направленного перебора: метод динамического программирования и метод ветвей и границ [21].

При использовании механизмов вывода, реализованных в ЭС, таких, как правила продукций, статистические методы принятия решений (байесовский подход, метод

Демпстера — Шейфера) и т. п., основной акцент переносится на методические вопросы, связанные с представлением знаний о диагностике в рамках данного механизма. Кроме того, большое внимание уделяется способам работы с неопределенностью (фактор уверенности, вероятность).

1.2.1. Сравнительный анализ характеристик существующих интеллектуальных систем тестирования

Анализ существующих тестирующих систем, использующих сеть Интернет с целью выявления уровня их интеллектуализации, осуществлялся по следующим критериям:

- 1) цели тестирования;
- 2) виды тестовых заданий;
- 3) выбор последовательности тестовых заданий;
- 4) критерии завершения тестирования;
- 5) способ информирования тестируемого;
- 6) вид представления результата теста;
- 7) характеристика предлагаемых вариантов ответов.

На основе проведенного анализа были получены следующие результаты:

– большинство тестирующих систем, существующих в настоящее время, основано на закрытой форме тестовых заданий с выбором одного варианта ответа из нескольких предложенных, имеющего определенный весовой балл. Это свойство редко позволяет учесть правильные ответы на основополагающие вопросы по предложенной теме, а также на неполные или частично правильные ответы;

– почти в каждой системе присутствует генератор тестовых заданий, который осуществляет их случайную выборку из общего числа имеющихся в базе. Иными словами последующее тестовое задание никак не учитывает правильность ответов, данных на предыдущие;

– во всех системах тестирование завершается лишь при прохождении всей строго определенной совокупности тестовых заданий. Это не позволяет сократить время тестирования за счет «идеальных» ответов на сложные тестовые задания, с помощью которых можно практически однозначно идентифицировать уровень знаний.

На основании сказанного можно сделать вывод, что практически ни в одной из доступных для анализа тестирующих систем (опубликованных и предлагаемых) не используются методы искусственного интеллекта [28].

В то же время применение интеллектуальных тестирующих систем позволяет:

- значительно улучшить качество тестирования;
- отслеживать успеваемость, как отдельного пользователя, так и всей группы;
- давать советы для более эффективного усвоения материала;
- подробно обосновывать человеку, прошедшему тестирование, полученную им оценку;
- сократить время прохождения теста за счет завершения процесса тестирования после наступления значимого события;
- оказывать помощь преподавателю по совершенствованию заданий теста и занятий в целом;
- выделять требования к качеству тестов (валидность, определенность, простота, однозначность, надежность).

Определим показатели качества тестов:

Валидность теста – это адекватность модели реальности. Различают содержательную и функциональную валидность: первая – это соответствие теста содержанию контролируемого учебного материала, вторая – соответствие теста оцениваемому уровню деятельности.

Выполнение требования определенности (общедоступности) теста необходимо не только для понимания каждым учеником того, что он должен выполнить, но и для исключения правильных ответов, отличающихся от эталона.

Требование простоты теста означает, что тест должен иметь одно задание одного уровня, т.е. не должен быть комплексным и состоять из нескольких заданий разного уровня. Необходимо отличать понятие “комплексный тест” от понятия “трудный тест”. Трудность теста принято характеризовать числом операций P , которое надо выполнить в тесте: $P < 3$ – первая группа трудности; $P = 3-10$ – вторая группа трудности. Не следует также смешивать понятия простоты-комплексности и легкости-трудности с понятием сложности.

Однозначность определяют как одинаковость оценки качества выполнения теста разными экспертами. Для выполнения этого требования тест должен иметь эталон. Для измерения степени правильности используют коэффициент $K_{\alpha} = P_1 / P_2$, где P_1 – количество правильно выполненных существенных операций в тесте или батарее тестов; P_2 – общее количество существенных операций в тесте или батарее тестов. Существенными считают те операции в тесте, которые выполняются на проверяемом уровне усвоения. Операции, принадлежащие к более низкому уровню в число

существенных не входят. При $K_\alpha \geq 0.7$ считают, что образовательная деятельность на данном уровне усвоена.

Понятие надежности тестирования определяют как вероятность правильного измерения величины K_α . Количественный показатель надежности $r \in [0, 1]$. Требование надежности заключается в обеспечении устойчивости результатов многократного тестирования одного и того же испытуемого. Надежность теста или батареи тестов растет с увеличением количества существенных операций P .

К безусловным достоинствам тестирующих программ можно отнести простоту их реализации. Создать тестирующую программу, даже использующую описанные выше усложненные схемы, может даже начинающий программист.

Достоинством является также и удобство использования таких систем пользователем. Выбирать правильный ответ из предложенных легче, чем самостоятельно получить результат. На это требуется меньше сил и времени.

Достоинством (хотя и не бесспорным) является то, что сама идея тестирования, т.е. предложение нескольких вариантов ответов, косвенно стимулирует пользователя анализировать различные решения, а как следствие – более глубоко исследовать поставленную задачу.

Тесты достаточно привлекательны, так как существенно сокращают время ответа на вопросы, а проверку тестов позволяют полностью автоматизировать.

Это открывает огромные перспективы использования тестирующих программ в среде дистанционного обучения, т.к. позволяют полностью реализовать весь процесс тестирования без участия преподавателя. Результаты тестов могут сохраняться автоматически и использоваться в дальнейшем. Этот принцип используют системы автоматического управления учебным процессом, создающие индивидуальные базы данных на каждого пользователя.

Однако имеются и недостатки. Основным недостатком является конечность баз данных. Для того, чтобы обеспечить успешное функционирование тестирующей программы в течение длительного времени, необходимо постоянно добавлять и расширять базы данных. Иначе постепенно программа перестанет выполнять свои функции.

Вторым недостатком является малая информативность ответа пользователя. Выбор правильного варианта ничего не говорит о том, как он был получен. Возможно, просто угадан.

Можно отнести к недостаткам тестирующих программ и то, что они накладывают

определенные ограничения и на сложность предлагаемых задач. Тест проходит за один сеанс и не допускает возможности отложить ответ, подумать, вернуться к тому же вопросу спустя день или два. Таким образом, естественно, что в тест можно включать только задачи, не требующие относительно больших усилий при решении.

И все же представляется весьма перспективным применение тестирующих программ в составе интеллектуального обучающего комплекса.

Также следует отметить роль интеллектуальных алгоритмов в работе тестирующих программ. Применение интеллектуальных тестирующих систем позволяет отслеживать успеваемость как отдельного обучаемого, так и всей группы; давать советы для более эффективного усвоения материала; оказывать помощь преподавателю по совершенствованию заданий теста и занятий в целом.

1.2.2. Структура интеллектуальной системы тестирования

Задача построения тестирующих программ как составной части интеллектуальной обучающей системы обусловлена следующими преимуществами тестирования как способа проверки знаний:

1. Простота решения исходной задачи построения системы.
2. Возможность дополнения тестирующей системы в процессе использования.
3. Достаточно простая схема практического использования.
4. Привлекательность для пользователя за счет времени и усилий, затрачиваемых на проверку знаний.

Однако представляется весьма перспективным применение тестирующих программ в составе интеллектуального обучающего комплекса.

Прежде всего, следует отметить роль интеллектуальных алгоритмов в работе тестирующих программ.

Использование интеллектуальных алгоритмов функционирования просматривается уже на этапе введения динамического изменения оценки вопросов и ответов, а также схемы прохождения теста. Очевидно, реализуются классические функции систем искусственного интеллекта – распознавание и идентификация ответа пользователя, принятие решения, формирование и реализация управляющего воздействия в виде оценки текущего и выбора следующего вопроса.

Еще одной реализацией интеллектуальных алгоритмов функционирования является возможность самоорганизации теста и его адаптации к уровню

подготовки пользователей.

Имеется в виду следующее: при многократном использовании теста, программа сама может перераспределять вопросы по уровню сложности, опираясь на данные, полученные при тестировании. Проще говоря, вопрос, на который по данным статистики было получено наименьшее число правильных ответов, автоматически переходит в разряд сложных, увеличивается его весовой коэффициент. Коэффициент вопроса, на который было получено наибольшее количество правильных ответов, напротив, снижается [32].

Это дает возможность построения адаптивных тестов, которые само корректируются под уровень пользователей. Например, две одинаковые в начале функционирования программы, применяясь в различных группах, через некоторое время будут существенно отличаться друг от друга.

В практической реализации этого принципа нет ничего сложного – программа просто ведет определенный файл протокола работы, из которого при запуске считываются данные. Однако, это – тоже один из интеллектуальных алгоритмов работы системы.

В составе интеллектуального обучающего комплекса система будет иметь автономную систему управления, решающую описанные выше задачи, но, кроме этого, она будет работать в составе системы управления всего комплекса. Автономная система управления будет участвовать в формировании и коррекции внутренней модели пользователя.

Таким образом, если будет образован один из контуров обратной связи интеллектуального обучающего комплекса – получая от главной системы управления текущую модель пользователя, то тестирующая программа будет давать на выходе скорректированную модель в соответствии с результатами прохождения теста (Рис.3).



Рис.3. Система управления интеллектуальной обучающей системы

В составе общей системы, возможно, существенно расширить круг задач, решаемых тестирующей программой.

Наиболее часто тестирующие программы используются для контрольного тестирования, однако такая ограниченность говорит о том, что как раз в этом качестве они не очень эффективны. Именно из-за малой информативности ответа пользователя, гораздо более перспективным представляется использование тестирующих программ для входного тестирования, самотестирования и промежуточного корректирования алгоритма обучения.

При таком использовании система оценки ответов пользователя в баллах теряет свой смысл. Тестирующая программа начинает функционировать как интеллектуальная экспертная система. Подобно экспертной системе, проводящей диагностику технических систем, тестирующая программа проводит своего рода диагностику пользователя, формируя внутреннюю модель обучающей системы.

Функционируя в составе интеллектуальной обучающей системы, тестирующие программы могут программно формировать схему дальнейшего обучения на основе полученных от пользователя ответов. Таким образом, в составе интеллектуальной обучающей системы, роль тестирующих программ состоит в обновлении модели

пользователя на основе использования баз данных задач.

Интерактивность при работе с тестирующей программой традиционно сводилась лишь к некоторому разнообразию способов обмена информации между системой и пользователем, а также к различной степени эргономичности интерфейса. Под ограниченностью интерактивности понимается невозможность пользователя влиять на исходные условия поставленной задачи. Идея допустить такую возможность вплотную приводит к созданию тестирующих программ с элементами искусственного интеллекта. В таких системах ответ пользователя влияет не только на тип следующего вопроса, но и на саму базу данных. Для этого локальная система управления тестирующей программы должна иметь дополнительную аналитическую подсистему.

Кроме того, тестирующие программы могут самостоятельно формировать и предлагаемые пользователю материалы в соответствии с качеством изучения курса.

1.3. Постановка задачи разработки и внедрения в учебный процесс интеллектуальной системы тестирования

Рассмотренные выше требования, предъявляемые к интеллектуальным обучающим системам по:

- мотивации состояний обучаемого;
- дифференциации требований и заданий к обучаемому;
- активизации развивающих интеллектуальных тестов;
- различным конфигурациям архитектуры с учетом типа, состава и назначения системы;
- функциональным особенностям различных подсистем всей системы;
- соблюдению выполнения качественных показателей тестов и др.

Все это выдвигает задачу разработки и внедрения в учебный процесс такой интеллектуальной системы тестирования, в которой будут учтены не только перечисленные требования, но и даны рекомендации по следующим вопросам:

- преемственность и простота реализации;
- реализация интеллектуальных алгоритмов динамически адаптирующихся к уровню подготовки пользователей и системы оценивания;
- разработка автоматизированной системы управления по формированию и коррекции внутренней модели пользователя;
- разработка тестирующей программы и ее апробация для входного

тестирования, самотестирования и промежуточного корректирования алгоритма обучения.

Выводы к 1 главе

Предложенная концепция тестирующей системы дает возможность использовать результаты оценки интеллектуального уровня пользователя и предлагает набор тестов, адаптированный к уровню подготовки тестируемого. При этом система перераспределяет вопросы по уровню сложности, опираясь на данные, полученные при тестировании.

Применение интеллектуальных алгоритмов позволяет также динамически изменять систему оценивания и схему прохождения теста, что существенно повышает качество и скорость тестирования.

Интеллектуальные системы тестирования могут использоваться как отдельно для проверки уже накопленных знаний, так и в составе обучающего комплекса для более эффективного усвоения приобретенных знаний.

ГЛАВА 2. ТЕХНОЛОГИЯ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ

2.1. Модели построения интеллектуальной системы тестирования

Определенное распространение получили попытки использовать в учебном процессе системы искусственного интеллекта. Именно применение методов искусственного интеллекта и инженерии знаний, поможет превозмочь субъективность и прямолинейность процесса тестирования, поднимая уровень оценивания знаний машинными системами. Дело в том, что применение традиционных методов компьютерного обучения и контроля (прямое тестирование, балльная система, и т.д.) неприемлемо: для того, чтобы оценить знания обучаемого, преподавателю приходится переработать огромное количество информации, и процесс взаимодействия оценивающего и оцениваемого не поддается строгой формализации, поэтому три основные алгоритмические структуры (последовательная, ветвление, цикл) при реализации машинных систем не смогут описать в полной мере данную предметную область. Т.е. осуществление автоматизированного контроля знаний, умений обучаемых, в первую очередь, включает решение проблемы определения совокупности требуемых качеств знаний, без которых критерии оценки знаний и способы определения уровня их усвоения выявить нельзя.

Основная цель контроля знаний и умений состоит в обнаружении достижений, успехов учащихся, в указании путей совершенствования, углубления знаний, умений, с тем, чтобы создавались условия для последующего включения учащихся в активную творческую деятельность. Эта цель в первую очередь связана с определением качества усвоения учащимися учебного материала – уровня овладения знаниями, умениями и навыками предусмотренных программой. Во – вторых, конкретизация основной цели контроля связана с обучением приемам взаимоконтроля и самоконтроля, формированием потребности в самоконтроле и взаимоконтроле. В - третьих эта цель предполагает воспитание у учащихся таких качеств личности, как ответственность за выполненную работу, проявление инициативы.

Интеллектуальное тестирование предполагает наличие модели знаний, модели самого процесса тестирования и модели оценивания. Так можно охарактеризовать, в общем, все разработки в этой области. Рассмотрим некоторые из них более подробно.

2.1.1. Модели распознавания образа уровня знаний

Традиционная система оценивания знаний обучаемых основана на лингвистических оценках, по которым устанавливается стипендия, производится учет успеваемости, проставляются записи в зачетных книжках за период обучения и др.

Вместе с тем, такая новая образовательная процедура как образовательное тестирование по альтернативному признаку предполагает оценивание уровня знаний в диапазоне от нуля до ста, что порождает проблему распознавания лингвистического образа знаний по результатам такого образовательного тестирования.

Под образом уровня знаний понимаются обучаемые, принадлежащие к множеству (группе), знания которых по “этalonу уровня знаний” отнесены к лингвистическим оценкам неудовлетворительно (**D**), удовлетворительно (**C**), хорошо (**B**), отлично (**A**).

Под распознаванием образа уровня знаний понимается процедура принятия решения о принадлежности конкретного обучаемого к одному из указанных образов на основании сравнения его образовательных достижений при тестировании с характеристиками образа.

При тестировании по альтернативному признаку используется закрытая форма теста, характеристиками которой являются: функция плотности распределения неправильных ответов $f(d)$, приемлемый уровень неправильных ответов q_0 , неприемлемый уровень неправильных ответов q_1 , риск заниженной оценки знаний a , риск завышенной оценки знаний b , функция оценивания знаний $f(Q)$, объем образовательной информации N , объем выборки заданий теста n и критерий принятия решений в виде предельного числа неправильных ответов K .

Перечисленные характеристики являются взаимозависимыми, но не обладающими достаточным свойством четкости. В условиях их нечеткости для распознавания образа уровня знаний обучаемых вполне допустимо для нормально реализованной образовательной услуги принять модель распределения неправильных ответов по закону редких случайных событий Пуассона и функцию оценивания уровня знаний сформировать по этому же закону.

Поскольку образовательная информация в банке заданий теста N в их выборке n представляется как статистическая совокупность, а задания теста обучаемому в компьютерном варианте всегда для выполнения выдаются последовательно, то для распознавания образа уровня знаний, возможно, воспользоваться последовательным критерием Вальда. При этом прием дополнительное принципиальное условие, что задания теста однородны по количеству образовательной информации по конкретной

учебной дисциплине, поскольку аналитических методов классификации заданий по мере их сложности или трудности пока не разработано.

Будем обозначать гипотезу о приемлемом уровне знаний H_0 , а гипотезу о неприемлемом уровне знаний H_1 . Пусть в результате последовательного поступления заданий теста в объеме n получены неправильные ответы $d_1, d_2, d_3, \dots, d_n$. При известной функции оценивания знаний по закону Пуассона последовательный критерий Вальда позволяет по выборке объемом n классифицировать обучаемых по уровню знаний на три подобраза по количеству областей принятия решений. Для того, чтобы иметь четыре образа необходимо произвести для каждой из трех областей повторное последовательное тестирование.

В предлагаемой процедуре рекомендуется использовать два способа распознавания образа уровня знаний: нормальный и усиленный. При этом задаются только четыре исходные характеристики теста q_1, q_0, a и b .

По нормальному способу по первой выборке заданий теста n_1 производится классификация обучаемых на три предварительные области (уровни): низкая, нормальная и высокая. По второй выборке заданий теста $n_2 = n_1$ или $n_2 < n_1$ для уровня низкий ужесточаются исходные характеристики q_0 и q_1 и обучаемые аттестуются по трем образам D, C и B. Для нормального уровня ужесточаются характеристики a и b и обучаемые аттестуются по трем образам C, B и A. Для высокого уровня тестирование осуществляется без изменения исходным q_0, q_1, a и b и обучаемые аттестуются на два образа B и A. К достоинствам нормального способа относится то, что обучаемые по второй выборке могут существенно улучшить свои образовательные достижения, что отвечает требованиям закона «О защите прав потребителей продукции и услуг».

По усиленному способу по первой выборке обучаемые классифицируются только на два уровня: низкий и высокий. По второй выборке для низкого уровня ужесточаются характеристики q_0 и q_1 и обучаемые аттестуются только на два образа D и C. По второй выборке для высокого уровня ужесточаются характеристики a и b и обучаемые аттестуются только на два образа B и A. Достоинством усиленного способа является более уверенное распознавание образа уровня знаний и поэтому его рекомендуется использовать в тех случаях, когда снижено доверие к реализуемой образовательной услуге.

Существенным отличием предлагаемых методов распознавания образа уровня знаний при тестировании от известных является заранее заданная погрешность распознавания, заложенная в рисках принятия решений a и b , использование наиболее

мощного критерия Вальда и достаточно простые и апробированные в статистическом приемочном контроле способы ужесточения планов тестирования, что позволяет использовать для решения практических задач распознавания образа уровня знаний международный стандарт ИСО 8423-91 «Статистические методы. Последовательные планы выборочного контроля по альтернативному признаку». Это способствует повышению достоверности компьютерного тестирования, что позволяет использовать предложенные теоретические разработки для обоснования методов оценивания знаний в предлагаемом едином экзамене до вузовского образования, в процессе реформирования системы образования.

2.1.2. Модель предметно - критериального составления тестов

В каждом курсе есть ключевые моменты, особенно важные темы, без знания которых невозможно усвоение более сложного материала в процессе учебы или которые будут необходимы в работе по специальности. На устном экзамене при личном контакте со студентом преподаватель обязательно оценивает понимание студентом этих тем. При автоматизированном тестировании можно учесть важность каких-либо разделов курса, увеличив долю вопросов по этим разделам в общем количестве вопросов. Но это не всегда удобно для составителя теста, потому что не всегда наиболее важные разделы содержат больше всего материала.

Предлагаемая методика предусматривает учет таких параметров, как степень важности и объем изучаемого материала в разделах курса.

При составлении теста преподаватель делит курс на темы T_1, T_2, \dots, T_k и оценивает степень важности S_i и объем изучаемого материала V_i по каждой теме T_i . Количество вопросов n_i по каждой теме T_i должно соответствовать (быть пропорционально) объему изучаемого материала V_i .

Минимальное количество вопросов n_i по каждой теме T_i определяется в соответствии с методикой с учетом параметра V_i .

Знания по каждому разделу курса оцениваются по пятибалльной (а фактически по четырехбалльной) системе. Оценке «отлично» (5) соответствует вероятность правильного ответа от p_3 до 1; оценке «хорошо» (4) соответствует вероятность правильного ответа от p_2 до p_3 ; оценке «удовлетворительно» (3) соответствует вероятность правильного ответа от p_1 до p_2 ; оценке «неудовлетворительно» (2) соответствует вероятность правильного ответа менее p_1 . Следует отметить, что вероятности p_1, p_2 и p_3 ($0 < p_1 \leq p_2 \leq p_3 < 1$) задаются

преподавателем с учетом структуры теста и могут быть изменены. Абсолютное количество (или доля) правильных ответов, достаточное для получения соответствующей оценки, определяется по специальной методике [16].

Итак, преподаватель:

- разбивает курс на темы (разделы) T_1, T_2, \dots, T_k ;
- определяет их объемы V_1, V_2, \dots, V_k и степень важности S_1, S_2, \dots, S_k ;
- определяет структуру теста – количество m вариантов ответов на каждый вопрос;
- задает p_1, p_2, p_3 – уровни знаний студента (или вероятности выбора правильного ответа), соответствующие оценкам: “2” – $0 \leq p < p_1$, “3” — $p_1 < p \leq p_2$, “4” — $p_2 < p \leq p_3$, “5” — $p_3 < p \leq 1$. p_1 должно быть заметно больше $1/m$ – вероятности выбора правильного ответа наугад.

После этого вычисляется минимальное количество вопросов n , необходимое для того, чтобы при заданных параметрах m, p_1, p_2, p_3 и заданном уровне значимости ϵ на основании испытания статистических гипотез можно было поставить оценку «5», «4», «3» или «2» за определенный раздел курса.

Минимальное количество вопросов n будет содержать тест по теме с минимальным значением $V_j = \min\{V_1, V_2, \dots, V_k\}$; $n_j = n$.

Минимальное количество вопросов по темам T_1, T_2, \dots, T_k определяется пропорционально их объемам, V_1, V_2, \dots, V_k .

По ответам студента вычисляется оценка O_i по каждой теме T_i ($1 \leq i \leq k$) как результат испытания статистических гипотез

При вычислении итоговой отметки за тест (курс) O учитывается степень важности S_i каждого раздела T_i . Получившаяся итоговая оценка O округляется до целых.

Следует отметить, что описанная выше методика позволяет давать студентам тест поэтапно, по мере изучения и усвоения материала отдельных разделов курса, и выводить итоговую оценку с учетом результатов промежуточного тестирования.

2.1.3. Модель Раша

Система тестирования на основе модели Раша обладает важными достоинствами, среди которых, прежде всего, необходимо отметить следующие.

1. Модель Раша превращает измерения, сделанные в дихотомических и порядковых шкалах в линейные измерения, в результате качественные данные анализируются с помощью количественных методов. Это позволяет использовать широкий спектр статистических процедур.

2. Оценка трудности тестовых заданий не зависит от выборки испытуемых, на которых была получена и, аналогично, оценка уровня знаний испытуемых не зависит от используемого набора тестовых заданий.

3. Пропуск данных для некоторых комбинаций (испытуемый — тестовое задание) не является критическим.

4. Сама система тестирования достаточно проста, по сравнению с другими аналогичными системами она характеризуется наименьшим числом параметров — только один параметр уровня знаний для каждого испытуемого и только один параметр трудности для каждого задания.

5. Модель Раша опирается на четкие и конструктивные понятия "трудность задания" и "уровень знаний". Так, одно задание считается более трудным, чем другое, если вероятность правильного ответа на первое задание меньше, чем на второе, независимо от того, кто их выполняет. Аналогично, более подготовленный студент имеет большую вероятность правильно ответить на все задания, чем менее подготовленный.

6. Благодаря простой структуре модели существуют удобные вычислительные процедуры для многоаспектной проверки адекватности модели: для всего набора тестовых результатов, для каждого испытуемого, для каждого задания и для каждого конкретного ответа.

7. "Остатки", получаемые при аппроксимации результатов тестирования моделью можно использовать для выделения различных типов испытуемых.

Однако, несмотря на 40-летний опыт применения этой системы тестирования за рубежом во многих областях знания, прежде всего в образовании, медицине и психологии, до сих пор продолжаются дискуссии об истинной ценности и эффективности системы тестирования на основе модели Раша. До сих пор существуют две крайние точки зрения на эту модель тестирования.

Наиболее убежденные сторонники модели Раша утверждают следующее: "Можно ли собрать или построить или сформулировать данные так, чтобы они соответствовали определению измерения (модели Раша)? Если нет, — то такие данные бесполезны".

Их наиболее последовательные оппоненты утверждают следующее: "Данные — это данные, а модель — это конструкция исследователя, которая подвержена ошибкам". Например, при построении регрессии, выбрасывая те или иные данные, можно получить любую зависимость, но мы тем самым ограничиваем реальный мир данных. Таким образом, мы создаем искусственную переменную, о которой мало что знаем.

Для практики одним из наиболее важных критериев является точность оценивания. Поэтому выбор темы в значительной степени обусловлен противоречивой информацией относительно точности системы тестирования на основе модели Раша. Кроме того, не удалось найти работы, в которых проведен всесторонний анализ точности модели Раша. В известных работах только даются те или иные общие рекомендации по использованию этих моделей [22].

Чем больше точность, тем лучше работает модель. В случае отсутствия ошибок измерения любая модель в смысле точности измерения работает идеально. Но на практике ошибки всегда есть и поэтому важно знать, насколько точные оценки позволяет получать та или иная модель.

На основе имитационного моделирования исследуются точность оценивания уровней знаний и трудностей заданий, а также число итераций, требуемых для вычисления этих оценок (методом наибольшего правдоподобия) в многофакторной ситуации в зависимости от:

- диапазона уровней знаний испытуемых;
- диапазона трудностей заданий;
- степени соответствия диапазонов уровней знаний испытуемых и трудностей заданий;
- числа испытуемых;
- числа заданий;
- степени соответствия данных модели;
- доли пропущенных данных.

Для статистической обработки результатов моделирования используется многофакторный дисперсионный анализ.

2.2. Технология интеллектуального тестирования

Контроль знаний учащихся является одним из основных элементов оценки качества образования. Педагогический контроль является важнейшим компонентом

педагогической системы и частью учебного процесса. До сих пор результатом педагогического контроля безоговорочно считается оценка успеваемости учащихся. Оценка определяет соответствие деятельности учащихся требованиям конкретной педагогической системы и всей системы образования.

Субъективность оценки знаний связана в определенной мере с недостаточной разработкой методов контроля системы знаний. Нередко оценка темы, курса или его частей происходит путем проверки отдельных, часто второстепенных элементов, усвоение которых может не отражать овладение всей системой формируемых знаний, умений, навыков. Качество и последовательность вопросов определяются каждым учителей интуитивно, и часто не лучшим образом.

Каждый из применяемых методов и форм проверки имеет свои преимущества и недостатки, свои ограничения. Кроме того, к недостаткам существующей практики проверки и оценки знаний следует отнести стихийность, нерациональное использование методов и форм, отсутствие дидактической целенаправленности, игнорирование учителем характерных особенностей материала предмета и условий работы в классе, отсутствие систематичности в ее проведении. Нельзя не отметить роль психологических факторов, общую и специальную подготовку учителя, его личные качества (принципиальность, чувство ответственности). Все это, так или иначе, влияет на результат проверки и оценки знаний. Именно педагогический субъективизм является главной причиной, по которой зачастую предпочтение отдают компьютерным и тестовым формам контроля с минимальным участием педагогов.

Помимо исключения субъективизма тестирование дает ряд существенных преимуществ перед традиционными формами контроля знаний:

1. сокращение времени прохождения теста;
2. сокращение трудозатрат преподавателя и высвобождения времени для творческой работы с обучающимися;
3. мгновенное получение результатов, исключая неадекватность реакции обучаемого (при текущем контроле знаний в частности) и повышение его мотивацию к обучению, как отметили в своем исследовании Г. Кречетников и Н. Черненко;
4. повышения качества образовательного процесса за счет своевременной корректировки учебной программы;
5. уменьшение субъективности оценки знаний.

Естественно, не все необходимые характеристики усвоения можно получить современными средствами тестирования. Такие, например, показатели, как умение

конкретизировать свой ответ примерами, знание фактов, умение связно, логически и доказательно выражать свои мысли, некоторые другие характеристики знаний, умений, навыков диагностировать тестированием практически невозможно, что и дает широкое поле деятельности для творческой работы преподавателя, а также для использования методов искусственного интеллекта.

Практика показывает, что если тестирование занимает более полутора часов, то испытуемые с неохотой соглашаются отвечать на вопросы теста. С другой стороны, с точки зрения теории, чем тест длиннее, тем он надежнее. Возникающее противоречие между теорией и практикой решается компромиссом в ту или в другую сторону, в зависимости от конкретного случая. Поэтому в настоящее время наибольший интерес представляет реализация технологии тестирования и разработка тестирующей системы, основанной на применении методов искусственного интеллекта, которая позволит решить ряд существенных недостатков современных тестирующих систем и повысить качество образования одновременно.

Возможности интеллектуального тестирования

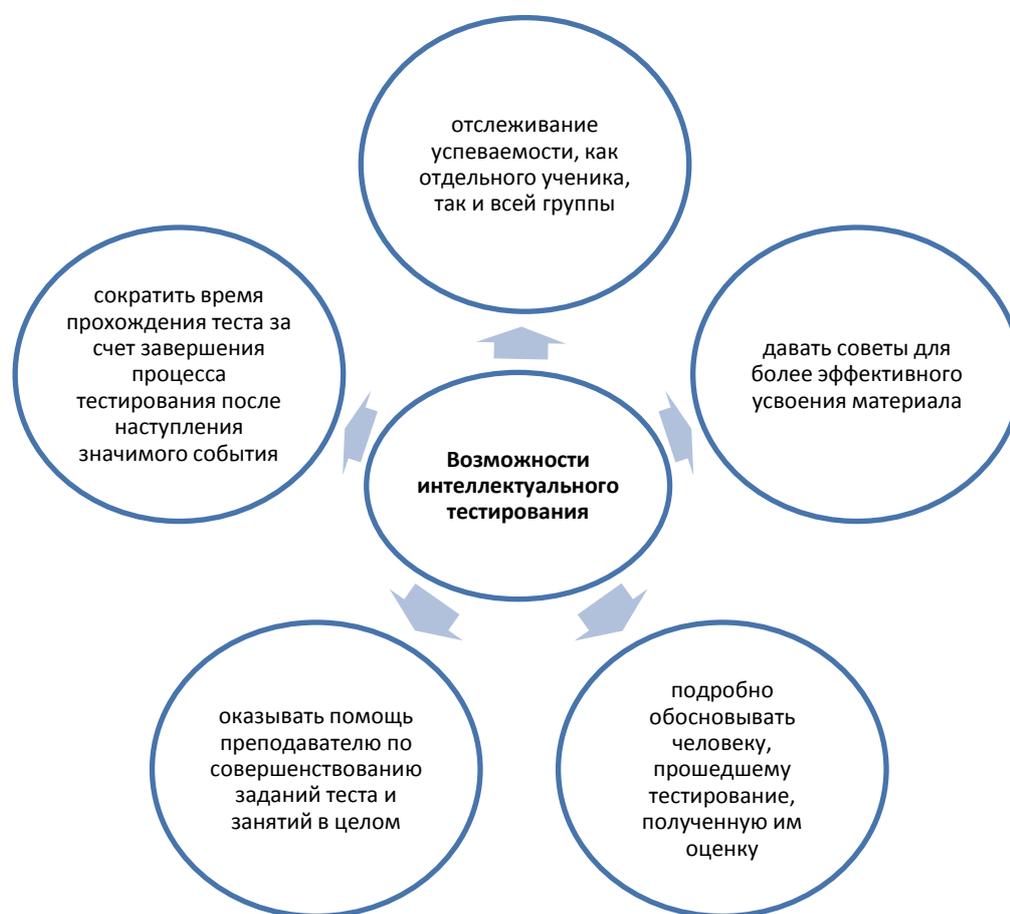


Рис. 4. Возможности интеллектуального тестирования

Системы такого рода должны быть направлены на облегчение эмпирической обработки и интерпретации результатов тестирования, снимая эту рутинную работу с преподавателя.

Введение интеллектуальных алгоритмов

Наиболее простые тесты имеют фиксированное количество стандартных вопросов, неизменную систему оценки полученных ответов, статический алгоритм построения последовательности теста.

Следующим шагом развития таких систем стало увеличение базы данных вопросов и ответов, а как следствие – разнообразие задаваемых программой вопросов. Как правило, эти вопросы выбирались случайным образом из базы данных. Подобный алгоритм позволил создать иллюзию уникальности предлагаемых тестов. Случайный выбор вопроса давал возможность формировать тесты с лавинообразным нарастанием количества вариантов.

Параллельно идет усложнение и системы оценки ответа пользователя. Первым шагом является статическое изменение весовых коэффициентов правильности ответа на тот или иной вопрос. База данных разбивается на группы по уровню сложности.

Вносятся изменения в систему управления тестом – она уже формирует тест, хотя и случайным образом, но так, чтобы вопросы были представлены в соответствии с уровнем сложности. Как правило, уровень сложности возрастает с номером вопроса. Прохождение теста уже допускается в произвольном порядке – к более ранним вопросам можно возвращаться. Это, вообще говоря, вносит в процесс тестирования определенную обучающую составляющую. Более поздние вопросы теста могут натолкнуть пользователя на правильный ответ на более ранний вопрос.

Второй шаг – динамическое изменение системы оценивания. Классический пример – предложение подсказки пользователю при неправильном ответе со снижением оценки за ответ.

Подсказок на один и тот же вопрос может быть и несколько, снижение баллов может быть пропорционально «ценности» подсказки. Пользователь может сам решать – использовать или не использовать ту или иную подсказку.

Традиционная компьютерная система контроля знаний предлагает каждому тестируемому ответить на одинаковое количество вопросов без учета качества ответов проверяемого на эти вопросы, после чего выполняется обработка результатов теста.

Эти тесты, как правило, включают перечень вопросов с вариантами ответов, среди которых находится правильный. Доля правильных ответов определяет итоговую оценку

уровня знаний тестируемого. Такие тесты можно назвать «жесткими», так как в них фактически используется двоичная система оценки правильности ответа на каждый вопрос (правильно или неправильно).

Основным недостатком «жестких» тестов является оценка неполных или неточных ответов как неправильных, т.е. приближенное, неполное знание ответа на вопрос квалифицируется как незнание ответа, а это далеко не всегда оправдано. В результате «жесткие» тесты недостаточно справедливо оценивают уровень знаний учащегося, занижают оценку. Поэтому выявляется необходимость работы с более «мягкими» тестами, позволяющими учитывать не только «знает» – «не знает», но и «частично знает» [8].

Создание тестирующих систем, работающих с «мягкими» тестами, позволяет решать и некоторые вопросы сопровождающего обучения при тестировании. Так как они позволяют вскрывать «анатомию» подготовки правильного ответа, что можно использовать при разборе ответа обучаемого и поиске именно того момента, который вызвал затруднение в его рассуждениях.

Динамически может меняться и схема прохождения теста. То есть программа может задавать вопросы определенного уровня сложности до тех пор, пока не будет получен правильный ответ (несколько правильных ответов), не переходя к более сложным вопросам. Это требует большого числа вопросов, а следовательно – больших баз данных.

Еще одной реализацией интеллектуальных алгоритмов функционирования является возможность самоорганизации теста и его адаптации к уровню подготовки пользователей.

Имеется в виду следующее: при многократном использовании теста программа сама может перераспределять вопросы по уровню сложности, опираясь на данные, полученные при тестировании. Проще говоря, вопрос, на который по данным статистики было получено наименьшее число правильных ответов, автоматически переходит в разряд сложных, увеличивается его весовой коэффициент. Коэффициент вопроса, на который было получено наибольшее количество правильных ответов, напротив снижается.

Это дает возможность построения адаптивных тестов, которые само корректируются под уровень пользователей. Две одинаковые в начале функционирования программы, применяясь в различных группах, через некоторое время будут существенно отличаться друг от друга.

В практической реализации этого принципа нет ничего сложного – программа просто ведет определенный файл протокола работы, из которого при запуске считываются данные. Однако, это – тоже один из интеллектуальных алгоритмов работы системы.

Интерактивность при работе с тестирующей программой традиционно сводилась лишь к некоторому разнообразию способов обмена информации между системой и пользователем, а также к различной степени эргономичности интерфейса. Под ограниченностью интерактивности понимается невозможность пользователя влиять на исходные условия поставленной задачи.

Идея допустить такую возможность вплотную приводит к созданию тестирующих программ с элементами искусственного интеллекта. В таких системах ответ пользователя влияет не только на тип следующего вопроса, но и на саму базу данных. Для этого локальная система управления тестирующей программы должна иметь дополнительную аналитическую подсистему.

2.3. Разработка концепции интеллектуальной системы тестирования на примере модели Раша

На основе результатов исследования была разработана концепция интеллектуального тестирования и создана интеллектуальная тестирующая система (ИТС), основанная на универсальном алгоритме вывода о знаниях, что позволяет использовать ее практически в любой предметной области.

На данный момент ИТС включает один сценарий работы и имеет следующий алгоритм. На первоначальной стадии система выдвигает 4 гипотезы: "испытуемый знает материал на 5", "испытуемый знает материал на 4", "испытуемый знает материал на 3", "испытуемый не знает материал". В этом первоначальном состоянии система дает одинаковую вероятностную оценку всем данным гипотезам.

Дальнейшая работа системы зависит от того, к какой гипотезе станет ближе тестируемый после ответа на последующие тестовые задания. Изменение уровней доверия между гипотезами происходит при помощи байесовского механизма. Таким образом, осуществляется адаптивность системы к уровню знаний испытуемого. При достижении критического уровня доверия к конкретной гипотезе человек, проходящий тестирование, получает оценку, а также информацию о недостаточно освоенном материале.

Работа ИТС в целом зависит, прежде всего, от сценария, заложенного изначально в систему автором комплекта тестовых заданий. Для применения ИТС в абсолютно любой области она должна обладать высоким уровнем универсальности, поэтому в основе системы может быть заложен как типовый сценарий, так и вновь созданный. Для этого планируется разработать специальный модуль, с помощью которого составитель теста

сможет самостоятельно задать сценарий работы системы, указав ряд таких свойств, как выбор последовательности тестового задания, механизм оценивания ответов, условие завершения тестирования, сохранение результатов тестирования. Таким образом, будет осуществляться поддержка деятельности преподавателя, в том числе интерактивная.

В отдельную задачу также можно выделить разработку модуля поддержки принятия решений, который будет помогать автору комплекта тестовых заданий составлять тест. Данный модуль будет направлен на анализ темы, материала, определяющего базовые и специальные знания в конкретной области, и проверку соответствия составленного теста данной теме и материалу [29].

Интеллектуальная тестирующая система позволит строить тест по одному из заданных сценариев, адаптироваться к уровню знаний испытуемого, сохранять результаты тестирования и использовать их для улучшения процесса обучения, а также давать рекомендации и советы для повышения уровня знаний испытуемых. Все это позволит повысить эффективность процесса обучения, практически исключив фактор субъективизма, и качество контроля знаний.

Применение методов искусственного интеллекта позволит преодолеть недостатки современных тестирующих систем, а также предоставить возможность для интерпретации таких навыков, как умение логически мыслить, конкретизировать ответы примерами, нереализованными в настоящее время.

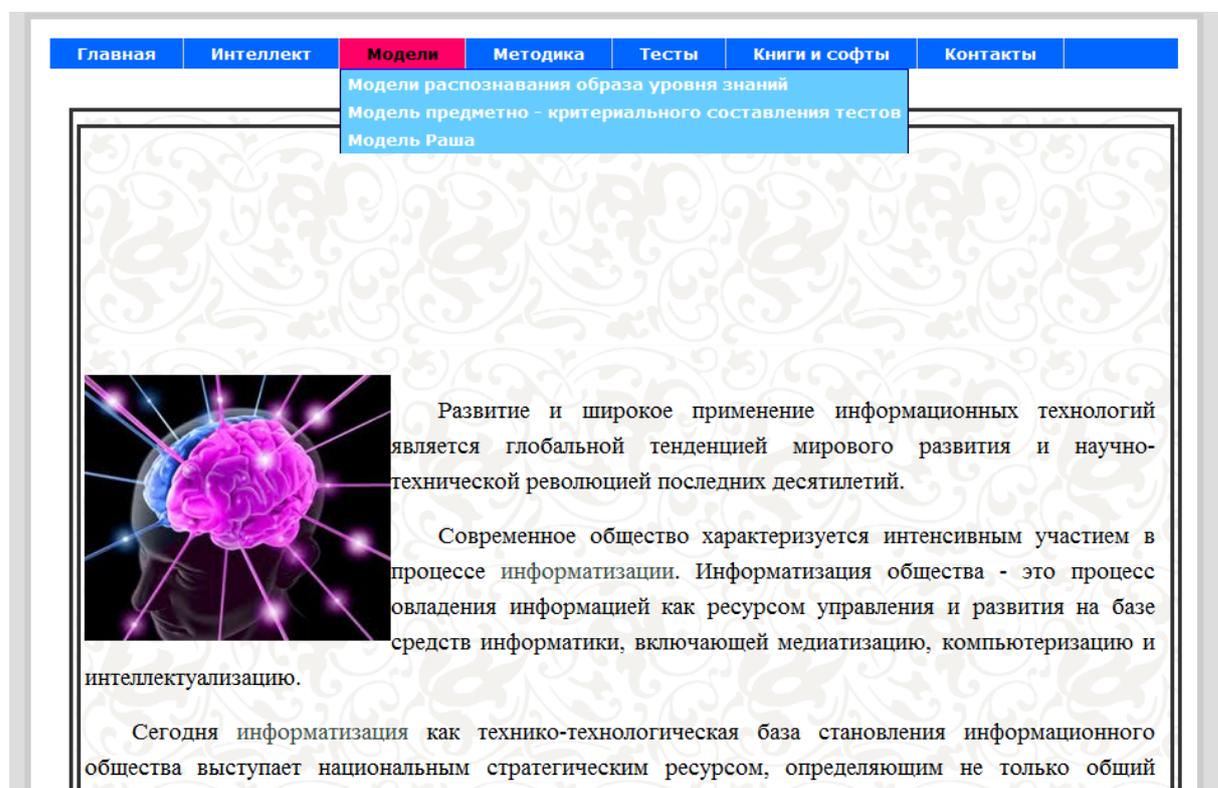


Рис. 5. Главная страница интеллектуальной системы тестирования

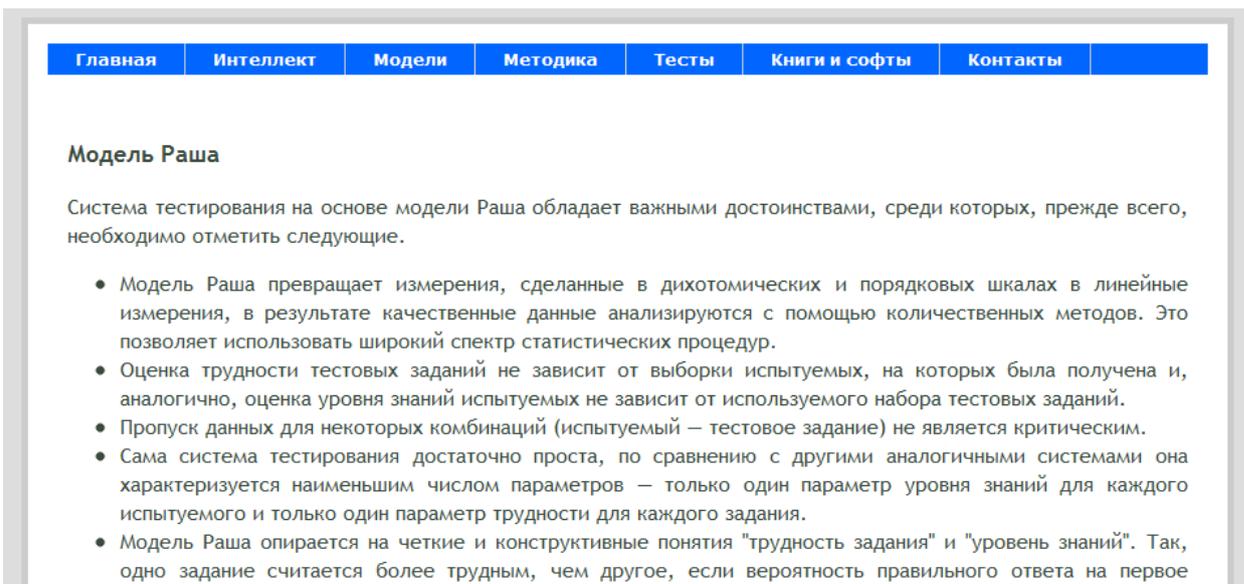


Рис. 6 . Описание модели Раша

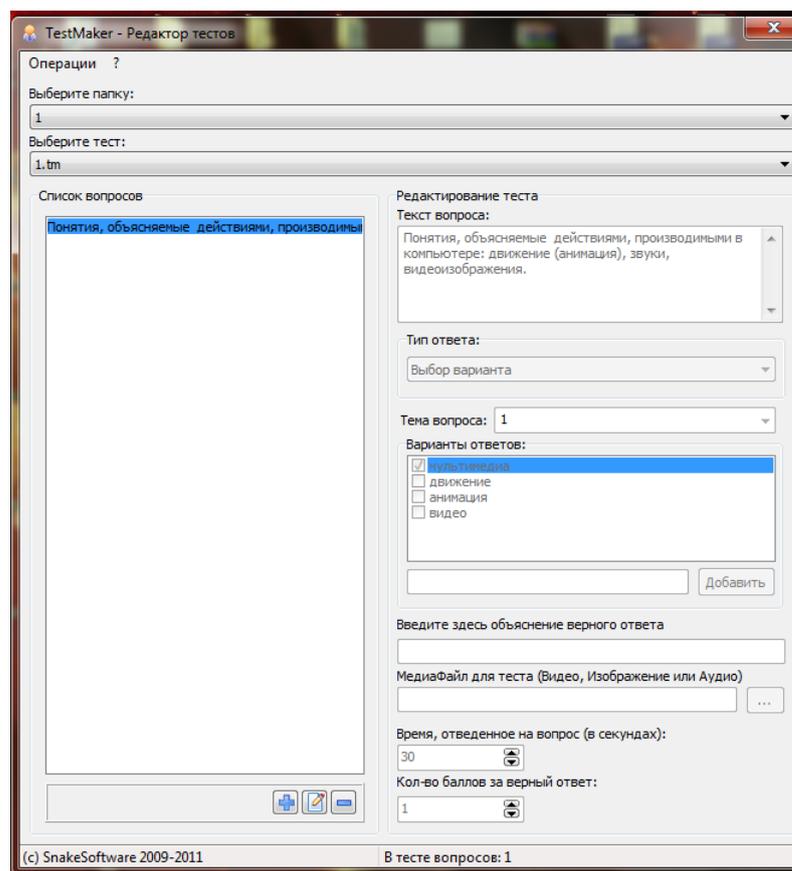
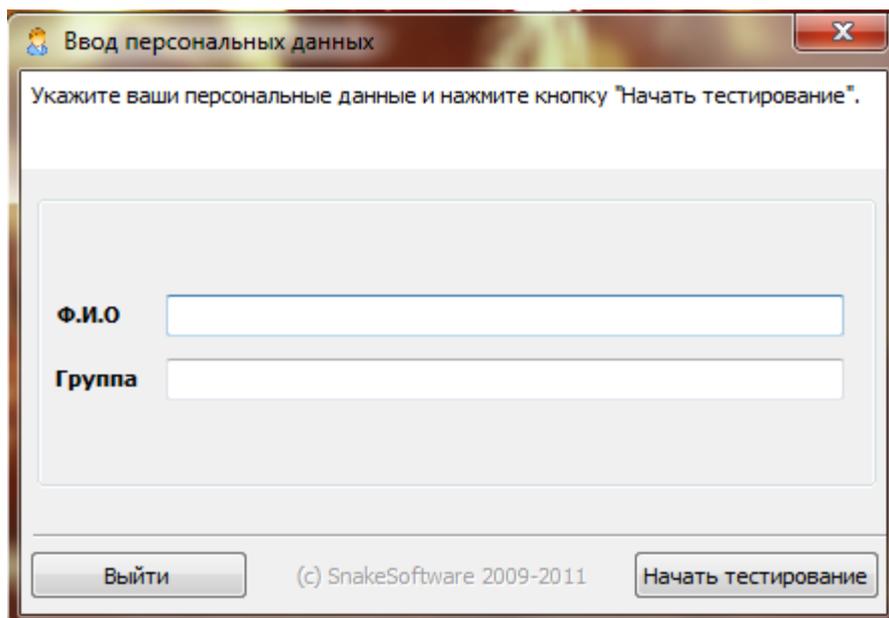


Рис. 7. Редактор тестов



Ввод персональных данных

Укажите ваши персональные данные и нажмите кнопку "Начать тестирование".

Ф.И.О

Группа

Выйти (c) SnakeSoftware 2009-2011 Начать тестирование

Рис. 8 . Окно ввода персональных данных

Выводы ко 2 главе

Во второй главе были проанализированы существующие методы и модели, позволяющие построить интеллектуальную систему тестирования. Сформулированы требования к построению модели интеллектуальной системы тестирования учитывающая: качество учебного материала, различные формы контроля знаний, необходимость выработки мотивационных составляющих.

Модель интеллектуальной системы тестирования построена как интегральная модель, базирующаяся на достоинствах таких моделей как: модель распознавания образа уровня знаний; модель предметно-критериального составления тестов; модель Раша.

Разработана технология интеллектуального тестирования с использованием интеллектуальных алгоритмов.

Разработана концепция интеллектуальной системы тестирования по сценарию модели Раша.

ГЛАВА 3. АПРОБАЦИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ В УЧЕБНОМ ПРОЦЕССЕ

3.1. Проведение педагогического эксперимента

Апробация является важным этапом в создании любого программного продукта. Во время апробации учитываются все ошибки и погрешности, допущенные разработчиком.

Эксперимент по апробации технологии интеллектуальной системы тестирования по курсу «Мультимедийные технологии в образовании» проводился в Ташкентском университете информационных технологий на факультете «Профессиональное образование» со студентами групп 253-09 и 433-09.

Апробация в данном случае является частью педагогического эксперимента. Педагогический эксперимент – сложный комплексный метод, включающий в себя целенаправленные наблюдения, беседы с испытуемыми, анкетирование, изучение творческих работ, социометрические измерения, организацию определенных видов групповой и индивидуальной деятельности [4].

Цель педагогического эксперимента – экспериментально проверить влияние выделенных условий на эффективность и качество обучения студентов с использованием технологии интеллектуальной системы тестирования по курсу «Мультимедийные технологии в образовании».

Выделим ряд задач, решаемых нами в ходе педагогического эксперимента:

1) Определить уровень усвоения знаний студентов по дисциплине «Мультимедийные технологии в образовании».

2) Экспериментально проверить влияние выделенных педагогических условий на качество усвоения знаний студентов по дисциплине «Мультимедийные технологии в образовании», с помощью технологии интеллектуальной системы тестирования.

3) Опираясь на результаты экспериментального исследования, доказать преимущество применения интеллектуального тестирования перед традиционным.

На эффективность применения экспериментального метода к исследованию педагогических явлений существенное влияние оказывает система общенаучных и конкретно-научных принципов, которая отражает общие требования к организации и проведению экспериментальной работы.

Достоверность получаемых в эксперименте результатов во многом зависит от условий, в которых он проводился, поскольку они могут оказывать прямое или косвенное

влияние на состояние или деятельность исследуемого педагогического объекта и, тем самым, выступать в качестве неконтролируемых экспериментальных переменных.

Условия организации и проведения экспериментальной работы:

Экспериментальная работа осуществлялась в естественных условиях, в ходе учебного процесса. В педагогическом эксперименте приняли участие студенты 253-09 и 433-09 групп. Количество испытуемых составляло 17 человек.

В рамках исследования был проведен параллельный эксперимент. При проведении параллельного эксперимента необходимо иметь как минимум две группы: контрольную и экспериментальную. Мы выделили две контрольные группы (II подгруппа 253-09 группы – контрольная группа, II подгруппа 433-09 группы – контрольная группа) и две экспериментальные группы (I подгруппа 253-09 группы – экспериментальная группа, I подгруппа 433-09 группы – экспериментальная группа). После проведения педагогического эксперимента сравнивалось состояние контрольных и экспериментальных групп.

При изучении дисциплины «Мультимедийные технологии в образовании», в контрольных группах проверка уровня усвоения знаний осуществлялась в рамках традиционного подхода – обычное тестирование, в экспериментальных группах проверка уровня усвоения знаний проводилась в рамках вводимой инновации – технология интеллектуальной системы тестирования.

Экспериментальная работа проводилась в три этапа:

Подготовительный этап эксперимента состоял в выявлении критериев и показателей, которые позволяют получить объективную информацию об эффективности экспериментальной работы.

Основной этап заключался в экспериментальном апробировании интеллектуальной системы тестирования по курсу «Мультимедийные технологии в образовании».

Заключительный этап состоял в следующем:

- 1) Провести анализ и интерпретацию экспериментальных данных.
- 2) На основе проведенного анализа, сформулировать выводы относительно результатов эксперимента.
- 3) Опираясь на результаты экспериментального исследования, доказать необходимость использования модуля контроля знаний по курсу «Мультимедийные технологии в образовании».

Определим следующие уровни усвоения знаний студентов с использованием разработанной технологии интеллектуальной системы тестирования по курсу «Мультимедийные технологии в образовании».

Низкий уровень усвоения знаний. Данный уровень характеризуется низкой успеваемостью студента и низкой активностью к познавательной деятельности (количество баллов от 0 до 25).

Средний уровень усвоения знаний. Данный уровень определяет среднюю успеваемость студентов (количество баллов от 25 до 75).

Высокий уровень усвоения знаний. Этот уровень характеризуется высокой успеваемостью студентов и качеством усвоения знаний, интересом студентов к предмету (количество баллов от 75 до 100).

Обозначим уровень усвоения знаний студентов следующим образом:

- 0 – низкий уровень усвоения знаний;
- 1 – средний уровень усвоения знаний;
- 2 – высокий уровень усвоения знаний.

3.2. Анализ результатов и эффективности интеллектуальной системы тестирования

Педагогический тест, как средство измерения учебных достижений, может дать достоверный результат только в случае его корректного применения. Корректность применения теста – это многоаспектное понятие, включающее в себя вопросы конструирования и дизайна теста, вопросы разработки и применения тестов и, разумеется, интерпретации результатов тестирования. В данной работе основное внимание уделено вопросам корректности интерпретации результатов педагогического тестирования, проводимого на основе модели Г. Раша. Анализ результатов обычно проводится на основе классической теории тестов или на основе Item Response Theory.

После выполнения работ по созданию теста и сбора данных на репрезентативной выборке испытуемых, производится интерпретация результатов [17].

IRT в настоящий момент является общепризнанной теорией. В качестве латентных параметров модели выступают как характеристики тестируемых, так и самого теста. Ю.Нейман и В.Хлебников делают вывод, что «...уникальность моделей семейства Г. Раша состоит в том, что они задают определенный механизм преобразования формальных наблюдений за исходом событий в объективные измерения на метрической шкале

латентных стимулов этих событий». Это очень важно, так как недостаточно глубокое осознание этого факта, может приводить к тому, что положения педагогических измерений могут критически восприниматься специалистами в области точных наук

Таким образом, несоответствие эмпирических данных модели Раша означает, что, например, имеются неточности в формулировке заданий, были нарушения в процедуре тестирования и т.д. Как отмечает В.Аванесов, в литературе можно встретить немало критики по поводу неприменимости модели Раша к множеству «тестов», и поэтому ведется поиск других моделей, более адекватных полученным результатам. Но здесь есть один очень важный вопрос. В теории Г. Раша никогда не ставилась задача адекватного описания данных. Напротив, это пример другой философии измерения - *model based measurement*, где утверждается противоположное – не модель должна соответствовать эмпирическим данным, а данные должны соответствовать модели. Об этом можно спорить, но в соответствии с философией Rasch шкалу (педагогический тест) образуют только те задания, которые отвечают данной модели измерения. Все остальные в тест не включаются.

В качестве критериев оценки результатов тестирования были избраны:

- среднее арифметическое значение коэффициента усвоения знаний;
- среднее квадратичное отклонение значения коэффициента усвоения;
- скорость выполнения;
- коэффициент объема усвоенных знаний;
- значение коэффициента эффективности процесса обучения,
- гистограммы значений коэффициента усвоения знаний.

Результаты выполнения контрольных заданий студентами в ходе тестирования по окончании самостоятельной подготовки к аудиторным занятиям фиксировались компьютером во внешней памяти - на магнитном носителе информации, в виде протокола выполнения тестовых заданий каждым студентом. В протоколе фиксировались фамилия студента, номер группы, номера учебных элементов, по которым студенту предлагались тестовые задания, номер теста из набора тестов по данному учебному элементу, количество существенных операций по каждому предложенному тесту, правильность ответа студента на каждый предложенный вопрос. Преподавателем, ведущим экспериментальное занятие, фиксировалось время выполнения студентами заданий (в экспериментальной группе - время работы с обучающей программой и время работы с тестовой программой, в контрольной группе - время работы с тестовой программой).

Обработка результатов эксперимента производилась следующим образом. Прежде всего рассчитывался коэффициент усвоения знаний каждым студентом. Расчет коэффициента проводился по формуле, предложенной В.П. Беспалько:

$$K = a/p$$

где a - количество правильно выполненных в ходе тестирования существенных операций,

p - количество существенных операций во всех предложенных студенту тестах.

Множество полученных значений коэффициента усвоения обрабатывалось с целью получения значений:

а) среднего арифметического значения коэффициента усвоения

$$K = \frac{\sum K_i}{n},$$

где K - среднее арифметическое значение K_i , K_i - элемент множества значений коэффициента усвоения, n - количество студентов, выполнивших тестовое задание;

б) среднеквадратичного отклонения значения коэффициента усвоения

$$\sigma = \sqrt{(\sum (K - K_i)^2) / (n - 1)},$$

где σ - среднеквадратичное отклонение значения коэффициента усвоения. Скорость выполнения контрольных заданий вычислялась по формуле:

$$K_{\text{скорости}} = K_{\text{пр}} / t_{\text{вып}},$$

где $K_{\text{пр}}$ - количество правильно выполненных заданий;

$t_{\text{вып}}$ - время, затраченное на их выполнение.

Коэффициент объема усвоенных знаний рассчитывался по формуле:

$$K_{\text{объема}} = N_{\text{усв}} / K_{\text{эл}}$$

где $N_{\text{усв}}$ - количество усвоенных элементов знаний, $K_{\text{эл}}$ - количество элементов знаний в эталоне.

Коэффициент эффективности процесса обучения рассчитывался по формуле:

$$K_{\text{эф}} = \frac{\sum m \cdot t}{M \cdot T},$$

где m - число студентов, t - продолжительность этапа занятия, M - общее число студентов в группе; T - продолжительность всего занятия.

Во время экспериментального занятия студенты должны были с помощью обучающей компьютерной программы подготовиться к предстоящему аудиторному занятию (на первом экспериментальном занятии - к практическому, на втором - к

лабораторному) и оценить результаты своей работы с помощью интеллектуальной системы тестирования.

Тема экспериментального занятия - "Мультимедиа". Первое экспериментальное занятие проводилось после лекции по данной теме, но перед практическим занятием.

Второе - после практического, но перед лабораторным занятием по данной теме

В качестве контрольной была выбрана группа 433-09 того же факультета.

Кроме того, расписание занятий этих групп позволяло провести контрольную проверку качества подготовки студентов группы 433-09 в тех же условиях, что и в группе 253-09. Студенты контрольной группы должны были изучить материал по конспекту лекций и по учебному пособию, рекомендованному ведущим лектором, с помощью специально разработанной системы заданий.

Экспериментальное занятие на первом этапе дало следующие результаты. В экспериментальном занятии приняло участие 3 студентов группы 433-09.

Средняя длительность работы с обучающей программой на первом экспериментальном занятии составила 35 минут, средняя длительность работы с тестирующей программой - 20 минут. Результаты тестирования экспериментальной группы даны в таблице 3.1.

Таблица 3.1

**Результаты тестирования студентов группы 253-09
(экспериментальная группа, 1- этап эксперимента)**

	Ф. И. О.	K_1	K_1	$K_1 - K_1$	σ
1	Абдулазизов Э.	0,734		0,004	
2	Реимбоев Ф.	0,66	0,73	0,07	0,05
3	Худжагелдиева Л.	0,816		0,086	
4	Югай Э.	0,816		0,086	
5	Абдуллаева Р.	0,66		0,07	
6	Жураева Х.	0,722		0,008	

Обработка данных, отображенных в таблице 3.1. дала следующие значения среднего арифметического коэффициента усвоения и его среднеквадратичного отклонения: $K_1 = 0.73 \pm 0.05$.

Полагая распределение значений коэффициента усвоения нормальным, мы вправе ожидать, что в области значений коэффициента усвоения от 0.7 и выше окажется большая часть любой группы студентов.

Значение скорости выполнения контрольных заданий оказалось равным **Кскорости** = **0.518** (среднее количество правильно выполненных заданий 9.85, среднее время, затраченное на их выполнение 19 мин.).

Коэффициент объема усвоенных знаний оказался равным **Кобъема** = **0.792** (среднее количество правильно усвоенных элементов знаний 9.85, среднее количество элементов знаний в эталоне 12.43).

Значение коэффициента эффективности процесса обучения оказалось равным **Кэф=1**, т.к. на всех этапах экспериментального занятия студенты находятся в дидактических системах, гарантирующих усвоение на первом уровне.

Гистограмма значений коэффициента усвоения имеет вид, приведенный на рис.3.1. Высота левого столбца, определяемая количеством студентов, имеющих коэффициент усвоения менее 0.7, соответствует 3 студентам. Высота правого столбца, зависящая от количества студентов, имеющих коэффициент усвоения не менее 0.7, соответствует 11 студентам. Таким образом, количество студентов, сформировавших в достаточной степени свои знания 1 уровня в результате работы с обучающей программой, составило 73% от общего количества студентов в группе.

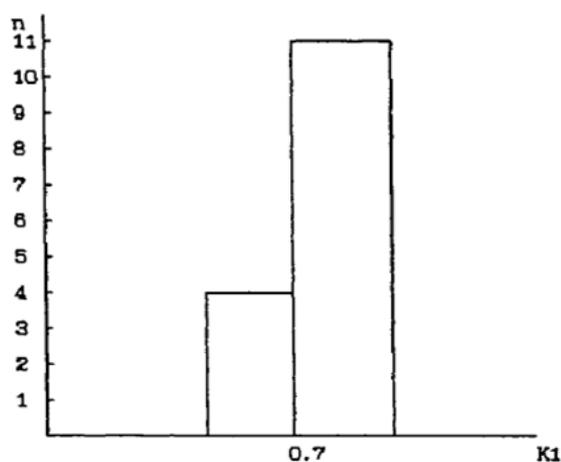


Рис.3.1. Распределение результатов тестирования студентов экспериментальной группы (первый этап эксперимента). n - количество студентов, K1 – значение коэффициента усвоения.

Контрольной группе 433-09 было предложено подготовиться к практическому занятию с помощью конспекта лекций и учебных пособий, традиционно используемых в курсе информационных технологий в коллежах. Как уже отмечалось, студенты пользовались специально разработанной системой заданий, посредством которой осуществлялось управление познавательной деятельностью студентов в ходе самостоятельной подготовки.

Для оценки результатов подготовки студенты этой группы должны были явиться в свободное от занятий время в компьютерный класс. На занятие явились 4 студентов.

В ходе контрольного занятия присутствующий на нем преподаватель фиксировал со слов студентов время, затраченное на самостоятельную подготовку к занятию и время, фактически затраченное студентами на выполнение тестовых заданий. Результаты выполнения тестовых заданий фиксировались компьютером на магнитном диске.

Среднее время, затраченное, по словам студентов, на самостоятельную подготовку к практическому занятию, составило 42 минуты. Среднее время выполнения тестовых заданий составило 35 минут.

Результаты тестирования контрольной группы даны в таблице 3.2

Таблица 3.2

**Результаты тестирования студентов группы 433-09
(контрольная группа 1-этап эксперимента)**

	Ф. И. О.	K_1	K_1	$K_1 - K_1$	σ
1	Абдукаримов У.	0,68	0,61	0,07	0,10
2	Абдуллаева С.	0,547		0,063	
3	Носиров Б.	0,468		0,142	

Обработка данных, содержащихся в таблице дала следующие значения среднего арифметического и среднеквадратичного отклонения значения коэффициента усвоения:

$$K_1 = 0.61 \pm 0.10.$$

Результаты обработки экспериментальных данных позволяют утверждать, что форма самостоятельной подготовки, используемая в контрольной группе, позволяет достичь требуемого качества лишь 15% студентов.

Значение скорости выполнения контрольных заданий оказалось равным:

$$K_{\text{скорости}} = 0.233$$

(среднее количество правильно выполненных заданий 8.14, среднее время, затраченное на их выполнение 35 минут).

Коэффициент объема усвоенных знаний оказался равным:

$$K_{\text{объема}} = 0.74$$

(среднее количество усвоенных элементов знаний 8.14, среднее количество элементов знаний в эталоне 11.0).

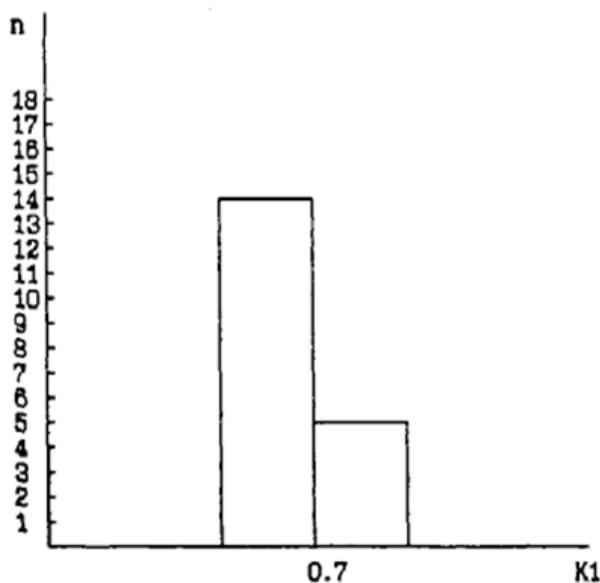


Рис. Распределение результатов тестирования студентов контрольной группы (первый этап эксперимента). n - количество студентов, K1 – значение коэффициента усвоения.

Значение коэффициента эффективности процесса обучения в контрольном занятии также оказалось равным 1, так как на всех этапах контрольного занятия студенты находились в дидактических системах, гарантирующих усвоение знаний на первом уровне.

Гистограмма значений коэффициента усвоения изображена на рисунке. Высота левого столбца гистограммы, пропорциональная количеству студентов, имеющих коэффициент усвоения менее 0.7, соответствует 14 студентам. Высота правого столбца, пропорциональная количеству студентов, имеющих коэффициент усвоения не менее 0.7, соответствует 5 студентам. Таким образом, традиционная методика подготовки позволила сформировать в достаточной степени знания у 26,3% студентов.

Сопоставление результатов самостоятельной подготовки в экспериментальной и контрольной группах позволяет отметить следующее.

Среднее значение коэффициента усвоения знаний в экспериментальной группе превысило требуемый (**K1 = 0.73**). В контрольной группе требуемый уровень достигнут не был (**K1 = 0.6**).

Разброс в качестве знаний (он характеризуется значением среднеквадратичного отклонения) в экспериментальной группе составил 0.05, в контрольной - 0.1, т.е. в два раза больше, чем в экспериментальной группе.

Экспериментальные данные позволяют предполагать, что в большей выборке, при использовании предлагаемой формы организации самостоятельной подготовки требуемое качество знаний будет сформировано у более чем 72% студентов, при традиционной - лишь у 15% студентов.

Значение скорости выполнения контрольных заданий в экспериментальной и контрольной группах отличаются более чем в два раза: $K_{\text{скорости}} = 0.518$ в экспериментальной и $K_{\text{скорости}} = 0.23$ в контрольной группе.

Значения коэффициента объема в экспериментальной и контрольной группах соотносятся как $K_{\text{объема}} = 0.792$ в экспериментальной и $K_{\text{объема}} = 0.74$ в контрольной.

Значения коэффициента эффективности процесса обучения в обеих группах оказались одинаковы.

Гистограммы результатов самостоятельной подготовки к практическому занятию наглядно показывают существенное улучшение качества подготовки в экспериментальной группе.

На втором экспериментальном занятии студенты как экспериментальной, так и контрольной группы должны были подготовиться к лабораторному занятию. Результатом самостоятельной подготовки должно быть завершение обучения на втором уровне усвоения (т.е. достижение $K_2 > 0.7$).

В экспериментальной группе на занятие были приглашены те студенты, которые приняли участие в первом экспериментальном занятии. Явились все 15 студентов.

Средняя длительность работы с обучающей программой второго уровня составила 32 минуты, средняя продолжительность работы с тестирующей программой - 15 минут.

Результаты тестирования студентов экспериментальной группы даны в таблице 4.3.

Обработка результатов тестирования студентов группы 270-08 дала следующее значение коэффициента усвоения:

$$K_2 = 0.73 \pm 0.05.$$

Свойства нормального распределения позволяют ожидать, что в области значений коэффициента усвоения от 0 до 0.6999 окажется не более 15% студентов, осуществлявших самостоятельную подготовку к лабораторной работе под управлением компьютера.

Таблица.3.3.

**Результаты тестирования группы 253-09
(второй этап эксперимента)**

Ф. И. О.	K_1	K_1	$K_1 - K_1$	σ

1	Турсунова У.	0.172		0.019	
2	Содикова З.	0.671	0.73	0.06	0.05
3	Пак Т.	0.81		0.049	
4	Эрмакова М.	0.793		0.062	
5	Исраилова Ф.	0,742		0,031	
6	Рахман-зода Дж.	0,658		0,258	

Скорость выполнения контрольных заданий оказалась равной

$$\mathbf{Кскорости = 0.415}$$

(среднее количество правильно выполненных заданий 8.71, среднее время на их выполнение 21 минута).

Коэффициент объема усвоенных знаний оказался равным:

$$\mathbf{Кобъема = 0.76}$$

(среднее количество усвоенных элементов знаний 8.71, среднее количество элементов знаний в эталоне 11.5).

Поскольку и на втором этапе все студенты находились в дидактических системах, гарантирующих усвоение знаний на втором уровне, коэффициент эффективности процесса обучения на втором этапе эксперимента оказался равным 1.

Из гистограммы значений коэффициента усвоения, изображенной на рисунке 4.3, видно, что количество студентов, не достигших в результате самостоятельной подготовки требуемого значения коэффициента усвоения, равно 3. Двенадцать студентов достигли значения коэффициента усвоения, позволяющего в дальнейшем самостоятельную познавательную деятельность. Другими словами, 80% студентов в результате работы с обучающей программой в достаточной степени сформировали знания 2 уровня [32].

Контрольная группа осуществляла самостоятельную подготовку к лабораторной работе в тех же условиях, что и на первом этапе эксперимента. В терминальный класс были приглашены студенты, принявшие участие в первом занятии. Явились все 19 студентов.

Среднее время, затраченное на самостоятельную подготовку в контрольной группе, оказалось равным 44 минуты (затраты времени оценивались со слов студентов). Среднее время выполнения тестовых заданий составило 38 минут.

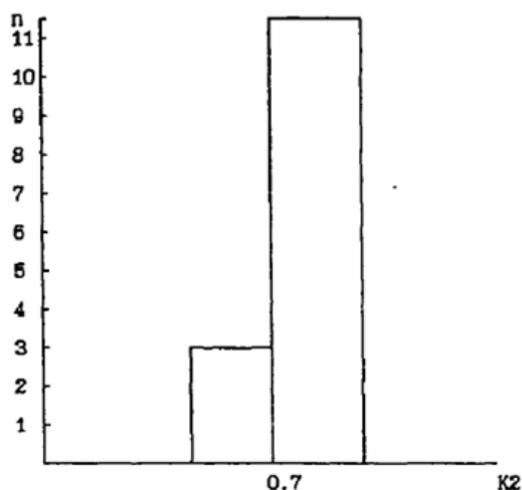


Рис. Распределение результатов тестирования студентов экспериментальной группы (второй этап эксперимента). n - количество студентов, К1 – значение коэффициента усвоения.

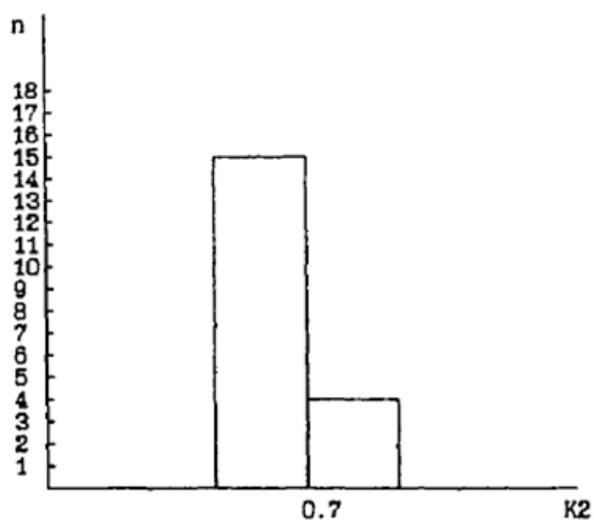


Рис. Распределение результатов тестирования студентов контрольной группы (второй этап эксперимента). n - количество студентов, К1 – значение коэффициента усвоения.

Результаты тестирования контрольной группы даны в таблице 3.4.

Среднее арифметическое значение коэффициента усвоения в контрольной группе оказалось равным

$$K2 = 0.58 \pm 0.09.$$

Полученный результат позволяет ожидать не более чем 10% студентов, в достаточной мере усвоивших требуемые знания.

Скорость выполнения контрольных заданий в контрольной группе

Кскорости= 0.172

(среднее количество правильно выполненных заданий 6.52, среднее время на их выполнение 38 минут).

Коэффициент объема усвоенных знаний оказался равным:

$$\mathbf{Кобъема = 0.61}$$

(среднее количество усвоенных элементов знаний 6.52, среднее количество элементов знаний в эталоне 10.70).

Целью самостоятельной подготовки к лабораторному занятию являлось формирование деятельности 2 уровня. При расчете коэффициента эффективности процесса обучения мы полагали

$$m \cdot t = 19 \cdot 38 = 722,$$

$$MT = 19 \cdot 82 = 1558.$$

Отсюда значение коэффициента эффективности процесса обучения

$$\mathbf{Кэф = 722/1558 = 0.4637.}$$

Из гистограммы, изображенной на рис.4.4, видно, что количество студентов, не достигших требуемого уровня усвоения в контрольной группе равно 15. Лишь четыре студента смогли с помощью традиционных средств самостоятельной подготовки достичь требуемого уровня усвоения.

Таким образом, 79% общего числа студентов не овладели в достаточной степени требуемыми знаниями.

Таблица.3.4.

**Результаты тестирования группы 433-09
(второй этап эксперимента)**

	Ф. И. О.	K_1	K_1	$K_1 - K_1$	σ
1	Хафизов Ш.	0.621		0.032	
2	Кокумбаева Е.	0.550	0.58	0.039	0.09

Выводы к 3 главе

Проведен педагогический эксперимент по апробации технологии интеллектуальной системы тестирования на примере курса «Мультимедийные технологии в образовании», где решены задачи: определения уровня усвоения знаний; выявление педагогических условий на качество усвоения знаний; сравнительной оценки результатов с использованием технологии интеллектуальной системы тестирования и с использованием традиционных методов.

Получены статистически подтвержденные показатели на контрольных группах обучаемых, которые продемонстрировали численно и графически достоинства интеллектуальной системы тестирования.

Заключение

В данной диссертационной работе были проанализированы существующие методы и модели, позволяющие построить интеллектуальную систему тестирования.

Основным научным результатом диссертационной работы является: предложена технологическая модель интеллектуальной системы тестирования, направленная на индивидуализацию подхода к каждому обучаемому, путем целевого выделения наиболее ориентированных частей системы; разработана технологическая структура интеллектуальной системы тестирования и ее возможные архитектурные преобразования в зависимости от поставленных целей.

Основные теоретические и практические результаты заключаются в разработке соответствующего алгоритмического и методического обеспечения, доступного для реализации в различных интеллектуальных системах тестирования.

Предложена концепция интеллектуальной системы тестирования, позволяющая использовать результаты оценки интеллектуального уровня для адаптации системы к уровню подготовки обучаемых.

Разработана концепция интеллектуальной системы тестирования по сценарию модели Раша.

Проведен педагогический эксперимент по апробации технологии интеллектуальной системы тестирования на примере курса «Мультимедийные технологии в образовании».

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Постановление Президента Республики Узбекистан от 02.06.2005 №ПП - 91 «О совершенствовании системы подготовки кадров в сфере информационных технологий».
2. Указ Президента Республики Узбекистан И. А. Каримова от 30.05.2002 «О дальнейшем развитии компьютеризации и внедрении информационно-коммуникационных технологий».
3. Аверченков, В.И. Система тестирования знаний с поддержкой ответов открытого типа на естественном языке / В.И. Аверченков, Е.А. Белов // Сборник научных трудов. Известия ОрелГТУ. Сер. Информационные системы и технологии. 2006. - № 1 (5). - С. 4-8.
4. Абуталиев Ф. Б. Возможности приложения теории нечетких множеств в системах поддержки принятия решений// «Инновацион-2000» Халкаро илимий-амалий анжумани. Илимий маколалар туплами. – Бухара, 2000. –С.271-273.
5. Абуталиев Ф. Б., Мухатдинов М. Я. Приложения теории нечетких множеств к проблемам принятия решений в управлении педагогическим процессом// Педагогик таълим. – Т., 2000. - №3. – С. 30-34.
6. Ахлиддинов Р. Ш., Саидов М. и др. Управление образованием в Узбекистане: проблемы, поиск, решения. - Т.: ЕС-ТАСИС, 1999.
7. Берестнева О.Г., Марухина О.В. Методы многомерного анализа данных в задачах оценки качества образования // Радиоэлектроника. Информатика. Управление. 2002. № 1. С. 15.26.
8. Бобомуродов О. Ж. Обработка информации в системах приобретения знаний. Издательство «ФАН» АНРУз, 2009г., 163с.
9. Верлан А. Ф., Чмыр И. А., Бобомуродлов О. Ж. Системы искусственного интеллекта. / уч. пособие. Самарканд, 2009г., 130с.
10. Головина Е.Ю. Чибизова И.В. О построении интеллектуальной обучающей системы. // Известия академии наук. Теории и системы управления. 1996. №5.
11. Дворяшина М.Д. Интеллектуальные способности и их влияние на успешность учебной деятельности студентов// Человек и общество. Проблемы повышения успеваемости и снижения отсева студентов/ Под. ред. В.Г. Лисовского, В.А. Сухина. -Л.: Изд-во ЛГУ, 1993.-Вып.ХХ.-С. 89-96.

12. Деревнина, А.Ю. Системы тестирования в электронных учебниках / А.Ю. Деревнина, В.А. Семикин // Информационные технологии. 2002. - №5. -с. 39-44.
13. Елисеева О.Е. Инструментальные средства проектирования интеллектуальных обучающих систем // «Высшая школа». 2008. -№1.
14. Жигалов И.Е., Таннинг Ж.Ф. Структура информационной обучающей системы с элементами искусственного интеллекта //Компьютерные технологии обработки и анализа данных. Под. Ред. С.С.Садыкова- Ташкент: Кибернетика.- 2000, с.79-85.
15. Карпенко, Д.С. Автоматизированная система мониторинга эффективности усвоения знаний и качества тестовых заданий / Д.С. Карпенко, О.М. Карпенко, Е.Н. Шлихунова // Инновации в образовании 2001. - №2 - С.69-86.
16. Колос В.В., Кудрявцева С.П., Сахно А.А. ПАЛЕВАС - система обучения основам логического программирования// Интеллектуализация компьютерных технологий обучения: Сб. науч. тр./НАН Украины. Ин-т кибернетики им. В.М. Глушкова. -Киев, 1993.-С. 50-53.
17. Ларин А.А. О концепции построения тестирующих программ [Электронный ресурс] / А.А. Ларин. – Режим доступа: alexlarin.narod.ru/Stats/testprog.htm
18. Майоров, А.Н. Теория и практика создания тестов для системы образования / А.Н. Майоров. М.: Народное образование, 2000. - 352с.
19. Мартынов Д.В., Смольникова И.А. Искусственный интеллект и образование //Тезисы научно-мет. конференции «Информационные технологии в образовании», Москва, 1999 -<http://ito.bitpro.ru/>
20. Марухина О.В. Алгоритмы обработки информации в задачах оценивания качества обучения студентов вуза на основе экспертно-статистических методов: Дис. канд. техн. наук: 05.13.01. .Томск, 2003. . 165 с.
21. Марухина О.В., Берестнева О.Г. Системный подход к оценке качества образования // Стандарты и качество. 2002. № 4. С. 35.36.
22. Моисеев, В.Б. Статистический подход к принятию решений по результатам тестирования для тестов открытой формы / В.Б. Моисеев и др. // Открытое образование 2001. - №1. - С. 51-57.
23. Минин М.Г. Диагностика качества знаний и компьютерные технологии обучения. Томск: Изд-во ТГПУ, 2000. .216 с.

24. Нейман Ю.М. Введение в теорию моделирования и параметризации педагогических тестов / Ю.М. Нейман, В.А. Хлебников. М., 2000. - 168 с. с табл. и илл. - стр. 6.
25. Нейман Ю.М. Вопросы точностных расчетов в теории моделирования параметризации педагогических тестов / Ю.М. Нейман // Труды Центра тестирования. М.: Прометей, 1999. Вып.2. - 49-54 с.
26. Нейман Ю.М. Основные модели современной теории тестирования / Ю.М. Нейман, ЕЛО. Карданова // Вопросы тестирования в образовании, 2003г. №7.
27. Петрушин В.А. Интеллектуальные обучающие системы: архитектура и методы реализации (обзор) // Известия Академии наук. Техническая кибернетика. -1993. №2
28. Саканян, Е.И. Использование системы АДОНИС для организации входного контроля знаний студентов при проведении практических занятий по технологии лекарственных форм / Е.И.Саканян, В.В.Клепиков- Режим доступа: <http://ito.bitpro.ru>
29. Система "Прометей" - Режим доступа: <http://www.jffometeus.ru/>64. "Прометей" - Режим доступа: <http://www.prometeus.ru/>65. "Аванта" - Режим доступа: <http://avanta.wsu.ru>
30. Суконщиков, А.А. Разработка экспертных и обучающих систем на основе принципов искусственного интеллекта / А.А. Суконщиков, А.Н. Швецов, А.П. Сергушичева // Сб. науч. трудов института в 2-х томах: Т.1. Вологда: ВоПИ, 1997. С. 32-35.
31. Хохлова О.С. Интеллектуальная тестирующая система / О.С. Хохлова // Новые информационные технологии: тезисы докладов XVI Международной школы-семинара. – М.: МИЭМ, 2008. – 297 с.
32. Чельшкова М.Б. Разработка педагогических тестов на основе современных математических моделей: Уч.пособие. М.: Исследовательский центр проблем качества подготовки специалистов, 1995;
33. Чельшкова М.Б. Теория и практика конструирования педагогических тестов: Учебное пособие. М.: Логос, 2002. .432 с.
34. Швецов, А.Н. Автоматизированные системы тестирования интеллекта/ А.Н.Швецов А.А.Сукошчиков // Матер, всерос. конф. "Математическое и компьютерное моделирование в науках о человеке и обществе". 28 июня- 2 июля 1999 г. Москва-Вологда.
35. Ахатова Р. Ю. Концептуальные подходы к созданию интеллектуальной системы обучения. // Республиканская научно-методическая конференция одаренных студентов,

магистрантов и аспирантов “Инновационные технологии и проблемы их теоретического и практического применения в учебном процессе”. – Ташкент. – 2011. – С. 40-42.

36. Ахатова Р. Ю. Теоретические основы современных технологий в образовании. // Республиканская научно-методическая конференция “Современные информационные технологии в телекоммуникации и связи”, посвященная 100-летию со дня рождения Исламова Анвара Исламовича. – Ташкент. – 2011. – С. 272-273.

37. <http://ito.edu.ru/2001/ito/VI/VI-0-1.html>;

38. <http://ito.edu.ru/2001/ito/VI/VI-0-10.html>

39. <http://ito.edu.ru/2001/ito/VI/VI-0-28.html>;

40. http://expro.ksu.ru/materials/ii_i_es/book.html