

МОСКОВСКИЙ АВТОМОБИЛЬНО-ДОРОЖНЫЙ
ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ (МАДИ)

04201351412

СВОБОДИН ВИТАЛИЙ ЮРЬЕВИЧ

**АВТОМАТИЗАЦИЯ И УПРАВЛЕНИЕ ПРОЦЕССОМ АТТЕСТАЦИИ
ПЕРСОНАЛА ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ НА ОСНОВЕ
КВАЛИФИКАЦИОННЫХ ХАРАКТЕРИСТИК И РЕЗУЛЬТАТОВ
АДАПТИВНОГО ТЕСТИРОВАНИЯ**

Специальность 05.13.06 – Автоматизация и управление технологическими
процессами и производствами (промышленность)

Диссертация
на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор технических наук,
профессор - Строганов В.Ю.

Москва – 2012

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
 1. АНАЛИЗ МЕТОДОВ И МОДЕЛЕЙ УПРАВЛЕНИЯ ПРОЦЕССОМ АТТЕСТАЦИИ ПЕРСОНАЛА.....	11
1.1. Проблемы кадрового обеспечения предприятий промышленности и транспортного комплекса	11
1.2. Педагогические принципы контроля и диагностики.....	13
1.2.1. Функции и методы контроля	14
1.2.2. Классификация педагогических тестов.....	19
1.3. Методики построения шкалы оценки уровня знаний	25
1.3.1. Принципы построения шкал в задачах тестового контроля	25
1.3.2. Формализованные модели классификации обученности	28
1.3.3. Унифицированная дидактическая классификация	36
1.4. Математическое моделирование процедур тестового контроля	39
1.4.1. Основные принципы классической теории тестового контроля	39
1.4.2. Формальные модели IRT-теории тестового контроля.....	41
1.4.3. Факторный, латентно-структурный и кластерный анализ в системе моделирования процедур тестового контроля.....	43
Выводы по главе 1	48
 2. ПОСТРОЕНИЕ МОДЕЛЕЙ КЛАССИФИКАЦИОННЫХ ХАРАКТЕРИСТИК И ОЦЕНОЧНЫХ ШКАЛ СИСТЕМЫ ТЕСТОВОГО КОНТРОЛЯ.....	50
2.1. Модель дифференцированной оценки сложности тестового задания.....	51
2.2. Параметризация логистических моделей Раца и Бирнбаума.....	53
2.3. Формирование матрицы ответов и первичных баллов.....	57
2.4. Формирование достаточных статистик оценки уровня знаний	66
2.5. Редукция матрицы ответов	68
2.6. Разработка моделей преобразования балльных шкал	70
Выводы по главе 2.....	74
 3. РАЗРАБОТКА ПРИНЦИПОВ КОНСТРУИРОВАНИЯ ТЕСТОВЫХ ЗАДАНИЙ И ФОРМАЛИЗАЦИЯ КОМПОНЕТОВ СИСТЕМЫ АДАПТИВНОГО ТЕСТОВОГО КОНТРОЛЯ.....	75
3.1. Построение формализованной модели тестового контроля	75
3.2. Принципы разработки тестовых заданий и формирование требований к инструментальной среде их конструирования	82
3.2.1. Разработка методических правил проектирования тестовых заданий ..	82
3.2.2. Формализованное описание механизмов формирования тестового задания	86
3.3. Формализованная марковская модель поведения адаптивного алгоритма	91
3.3.1. Построение марковской цепи процесса адаптивного тестового контроля	91

3.3.2. Сходимость стационарных вероятностей при увеличении дискретизации шкалы уровня сложности	94
3.3.3. Плотность распределения сложности тестовых заданий	97
3.3.4. Плотность распределения для функций регрессии с постоянной дисперсией и постоянным коэффициентом вариации	98
Выводы по главе 3	103

4. ПРОГРАММНО-МОДЕЛИРУЮЩИЙ КОМПЛЕКС И АПРОБАЦИЯ СИСТЕМЫ АДАПТИВНОГО ТЕСТОВОГО КОНТРОЛЯ

4.1. Проектирование структуры программного комплекса и интерфейса баз данных	104
4.1.1. Функции пользователей системы тестового контроля	104
4.1.2. Организация диалога в системе тестового контроля	105
4.2. Структура базы данных системы тестового контроля	109
4.3. Анализ результатов аттестации по возрастным группам сотрудников ..	113
4.3.1. Анализ зависимости по финансово-экономическому блоку	114
4.3.2. Анализ зависимости между результатами аттестации и возрастными категориями сотрудников по блокам	124
Выводы по главе 4	130

ЗАКЛЮЧЕНИЕ	131
-------------------------	------------

ЛИТЕРАТУРА	133
-------------------------	------------

ПРИЛОЖЕНИЕ. Акты внедрения результатов работы	146
--	------------

нет

Введение

Одной из основных задач системы подготовки и переподготовки персонала является оценка соответствия кадрового состава должностным обязанностям, что требует постоянного определения уровня знаний и умений сотрудников для оценки качества выполнения работ. Аттестация сотрудников промышленных предприятий представляет собой целенаправленный интерактивный процесс взаимодействия субъектов и объектов между собой, а также с соответствующими аппаратно-программными средствами. Система аттестации является одним из компонентов системы подготовки и переподготовки персонала, в которую также входят подсистемы определения целей, содержания, методов и средств обучения, организационно-правовых форм, а также учебно-материальная, финансово-экономическая, нормативно-правовая и маркетинговая подсистемы. Оперативный контроль уровня знаний и умений сотрудников может быть обеспечен лишь за счет всестороннего использования процедур компьютерного тестового контроля, обеспечивающих динамическую идентификацию уровня подготовленности персонала. Повышение эффективности компьютерного тестирования достигается за счет адаптивных алгоритмов, разработке которых в последнее время уделяется большое внимание. Однако в данном направлении отсутствуют работы по созданию комплексных методик, включающих методы, алгоритмы и программные компоненты формирования тестов вместе с процедурами анализа эффективности механизмов предъявления тестовых заданий, что и определяет актуальность настоящей диссертационной работы.

Предметом исследования являются система аттестации персонала, включающая методы компьютерного тестового контроля, методы обработки результатов, а также компоненты математического, информационного и программного обеспечения системы поддержки и принятия решений при классификации уровня знаний сотрудников промышленных предприятий.

Целью работы является повышение эффективности системы аттестации по критерию точности оценки квалификационных характеристик за счет разработки и использования методов, моделей, алгоритмов и программных средств поддержки компьютерного тестового контроля.

Для достижения данной цели в работе решаются следующие задачи:

- системный анализ методов и моделей организации процедур компьютерного тестового контроля в системе аттестации промышленных предприятий;
- разработка системы квалификационных характеристик, оценочной шкалы и формирование схемы групповой аттестации;
- разработка механизмов формирования тестовых заданий и требований к программной инструментальной среде;
- разработка и анализ эффективности методов адаптивного тестового контроля на основе создания интерактивной среды имитационного моделирования;
- разработка программно-моделирующего комплекса конструирования тестовых заданий и тестов.

При разработке формальных моделей компонентов системы в диссертации использовались методы общей теории систем и классический теоретико-множественный аппарат. Расчет вычислительных экспериментов выполнен в пакетах Statistica и MathCad. При разработке моделей компонентов системы тестового контроля и аттестации использовались методы математического программирования, теории случайных процессов, имитационного моделирования и др.

Структура работы соответствует списку перечисленных задач, содержит описание разработанных методов, методик и алгоритмов.

В первой главе диссертации проводится системный анализ процедур тестового контроля. Рассмотрены проблемы кадрового обеспечения промышленных предприятий и общие тенденции развития системы

непрерывного образования. Рассмотрены математические модели и методы моделирования процессов аттестации персонала.

При разработке процедур компьютерного тестового контроля основной проблемой является оценка эффективности предлагаемых процедур с точки зрения точности оценивания уровня знаний тестируемого. Адаптивный механизм предъявления тестовых заданий предполагает выбор каждого последующего задания из базы тестовых задания в зависимости от ответов на все предыдущие. В данном случае формирование алгоритма для описания механизмов предъявления дает большую свободу при построении новых методических принципов.

Предполагается, что база данных имитационной модели содержит все сформированные запросы по вариантам формирования структуры тестовых заданий. Результаты расчетов по желанию пользователя могут быть представлены как в табличной форме, так и в графической форме. Имеется возможность сравнительного анализа выбранных решений и оптимизации распределения при заданных ограничениях.

В общем случае оценивание уровня знаний по результатам ответов на тестовые задания в диссертации рассматривается как задача управления, где на каждом шаге процедуры выбирается очередное задание, обладающее определенной сложностью.

Во второй главе диссертации рассматриваются вопросы построения функциональных соотношений описания индивидуальных свойств обучаемого, которые базируются на моделях квалификационных характеристик.

С целью осуществления мониторинга уровня сформированности требуемых характеристик специалиста предлагается использовать комплексный показатель качества психолого-индивидуальных компетенций, где учитываются локальные коэффициенты сформированности профессиональных компетенций, число локальных коэффициентов.

Психолого-индивидуальные компетенции представляют собой способности, развиваемые в профессиональной деятельности под влиянием мотивации, которая может, как усиливать, так и ослаблять потенциальные задатки специалиста. В этой связи необходимо осуществлять постоянный мониторинг мотивационной направленности. Для определения критериев сформированности проводится тестирование выборки специалистов, успешных в своей деятельности. Для визуализации локальных показателей используется их представление в виде профилограммы, где отображаются коэффициент точности внимания; коэффициент продуктивности влияния; коэффициент распределения внимания; коэффициент объема памяти; коэффициент творческого мышления; коэффициент логического мышления; Применение такой модели специалиста на основе процессного подхода и методологии «развертывания функции качества» позволяет снизить уровень неопределенности идентификации и мониторинга рассматриваемых компетенций и повысить качество аттестации и переподготовки специалистов.

В третьей главе диссертации рассматриваются вопросы построения программно-моделирующей среды компьютерного тестового контроля в системе аттестации персонала. Первичным при формировании теста, является конструирование тестового задания. В работе предлагается формализованное описание методики конструирования тестовых заданий, с использованием разработанного с участием автора конструктора тестовых заданий, который имеет следующие функциональные компоненты: графический редактор формирования структуры задания; формирование схемы разбора ответа; параметризация выполнения задания; подсистема генерации Flash-файла тестового задания. Функции графического редактора используются для создания графического образа с интерактивными полями.

Круг задач тестового контроля, охватывающий основные аспекты моделирования (для методиста с целью оценки его эффективности) и непосредственного контроля, решается путем создания совокупности сценариев,

включающих: метод оценки вероятностей ошибочной классификации; ... (моделирующие алгоритмы); метод переоценки сложности тестовых заданий; язык формирования последовательности тестовых заданий; приложение компиляции языка; приложение разбора структуры гетерогенного теста; визуализация результатов тестового контроля и другие.

В четвертой главе диссертации рассматриваются вопросы построения программного комплекса автоматизации и моделирования процесса аттестации персонала. Система реализована в виде отдельного программного приложения. Программный комплекс разработан по открытому принципу, что позволяет наращивать его функциональные возможности, добавляя новые механизмы предъявления тестовых заданий в процессе эксплуатации системы.

Модуль построения сцены ответственен за визуализацию расположения объектов на сцене и задание свойств объектов. Содержит основные компоненты: инспектор компонент, инспектор свойств выделенного компонента, редактор сцены, палитра компонент. Эти компоненты взаимосвязаны друг с другом, таким образом, что изменения свойства объекта в одном компоненте (например, в инспекторе свойств) немедленно влечет обновление остальных компонент (например, изменения сразу отображаются в редакторе сцены). Модуль проектирования решения визуализирует граф решения. Позволяет задавать последовательность действий для правильного решения теста, указывать какие действия можно совершать параллельно. На основе данных графа модуль строит регулярное выражение, задающее класс правильных ответов на тест. Модуль скриптов теста позволяет добавить программные обработчики к тесту. Обработчики пишутся на языке программирования JavaScript. Обработчики динамически подключаются к тесту на этапе выполнения. Они позволяют реализовать любую функциональность теста, графическое задание которого было бы затруднительно (например, пересчет некоторых параметров при определенных действиях пользователей). Модуль упаковки теста позволяет

собрать XML документ теста вместе со всеми ресурсами и библиотеками в единую структуру, готовую для сохранения в базе данных и распространению конечным пользователям. Модуль проигрывания теста позволяет запустить тест в режиме проигрывания. Все действия пользователя автоматически записываются для дальнейшего анализа. Модуль обработки ответа пользователя проверяет правильность действий пользователя при прохождении теста и сравнивает его на совпадение с регулярным выражением, задающим правильный ответ теста.

В заключении представлены основные результаты работы.

В приложении приводятся акты внедрения результатов диссертационной работы.

Научную новизну работы составляют методы и модели организации аттестации персонала промышленных предприятий на базе адаптивных механизмов тестового контроля. **На защиту выносятся:**

- модели формирования балльных шкал и оценок квалификационных характеристик;
- комбинированный алгоритм вычисления сложности заданий и уровня подготовленности участников тестирования по схеме групповой аттестации;
- автоматная схема описания проигрывателя тестовых заданий с встроенным компилятором формирования последовательности тестовых заданий.

Обоснованность научных положений, рекомендаций и выводов определяется корректным использованием современных математических методов и моделей, предварительным статистическим анализом процессов обучения и тестового контроля ряда образовательных учреждений, согласованностью результатов аналитических и имитационных моделей процессов компьютерного тестового контроля. Достоверность положений и выводов диссертации подтверждена положительными результатами внедрения результатов работы в учебных центрах предприятий транспортного комплекса.

Научные результаты, полученные в диссертации, доведены до практического использования на транспортных предприятиях. Методы и алгоритмы, а также программные средства могут быть использованы также при реализации тестового контроля студентов высших учебных заведений. Внедрение результатов работы позволяет повысить качество и эффективность процессов разработки учебных программ и процедур тестового контроля.

Содержание отдельных разделов и диссертации в целом было доложено и получило одобрение:

- на Российских, межрегиональных и международных научно-технических конференциях, симпозиумах и семинарах (2009-2012гг.);
- на заседании кафедры АСУ МАДИ.

Совокупность научных положений, идей и практических результатов исследований в области автоматизации образовательного процесса составляет актуальное направление в области теоретических и практических методов и форм проведения тестового контроля и аттестаций сотрудников транспортных предприятий.

По результатам выполненных исследований опубликовано 7 печатных работ.

Диссертационная работа состоит из введения, четырех глав и заключения, опубликованных на 132 страницах машинописного текста, содержит 37 рисунков, 19 таблиц, список литературы из 123 наименований и приложения.

1. АНАЛИЗ МЕТОДОВ И МОДЕЛЕЙ УПРАВЛЕНИЯ ПРОЦЕССОМ АТТЕСТАЦИИ ПЕРСОНАЛА

В первой главе диссертации проводится системный анализ предметной области. Рассмотрены проблемы кадрового обеспечения промышленных предприятий и общие тенденции развития системы непрерывного образования. Рассмотрены математические модели и методы моделирования процессов компьютерного тестового контроля. Среди методов общесистемного моделирования особое внимание уделено методам принятия решений в условиях неопределенности. Проведен сравнительный анализ принципов конструирования тестовых заданий и построения тестов. Показано, что при обучении контроль рассматривается как важнейшее, относительно самостоятельное и заключительное звено в системе переподготовки кадров. От правильной организации контроля во многом зависит не только эффективность управления учебным процессом и качество подготовки специалистов, но и эффективность управления персоналом в целом.

1.1. Проблемы кадрового обеспечения предприятий промышленности и транспортного комплекса

Общеизвестно, что машиностроение и транспорт определяет перспективы развития индустрии в целом. В таких развитых странах, как США, Япония, Германия, на долю этой отрасли приходится до половины объема промышленной продукции.

Несоответствие квалификации сотрудников потребностям организации негативно сказывается на результатах ее деятельности. Развитие персонала – это важнейшее условие успешного развития любой компании. Организация профессионального обучения является одной из основных функций управления персоналом.

Ключевым моментом в управлении профессиональным развитием является определение потребностей организации в этой области. По существу речь идет о выявлении несоответствия между профессиональными знаниями и навыками, которыми должен обладать персонал и теми знаниями и навыками, которыми он обладает в действительности [7].

Важнейшим средством профессионального развития персонала является профессиональное обучение. В компании организация профессионального обучения должна представлять комплексный непрерывный процесс, включающий в себя несколько этапов. (рис. 1.1.)

Выполнение должностных обязанностей требует от сотрудников организации знания рабочих процедур и методов, выпускаемой продукции и оказываемых услуг, умения работать на установленном оборудовании и т.п.

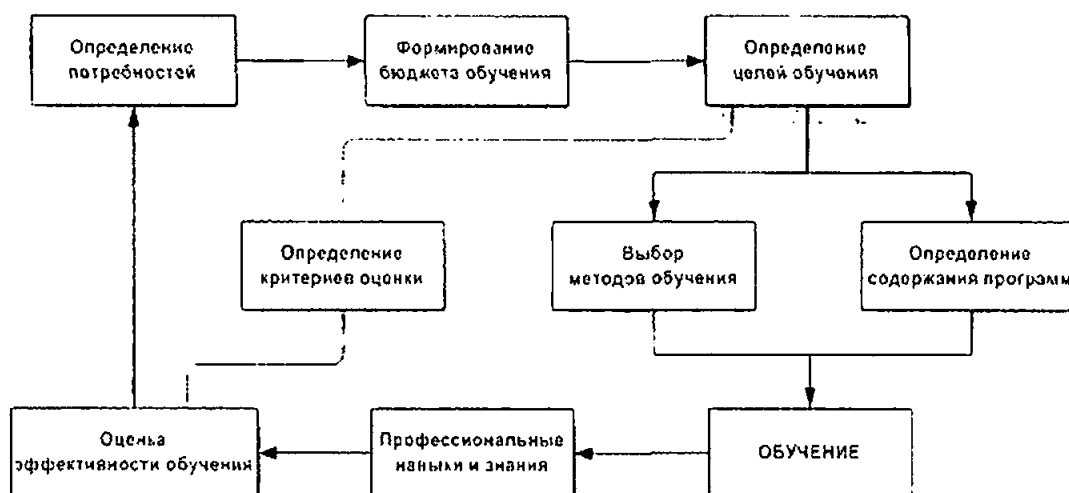


Рис. 1.1. Этапы организации профессионального обучения

Потребности, связанные с выполнением производственных обязанностей, могут определяться как на основе заявок руководителей подразделений и самих работников, так и путем проведения опросов руководителей и специалистов, анализа результатов работы подразделений организации [38]. Еще один источник информации о потребностях в

профессиональном обучении - индивидуальные планы развития, составленные по результатам аттестации сотрудников.

Индивидуализация обучения может быть обеспечена лишь за счет всестороннего использования процедур компьютерного тестового контроля, обеспечивающих динамическую идентификацию уровня подготовленности обучаемых. Повышение эффективности компьютерного тестирования достигается лишь за счет адаптивных методов. Однако в данном направлении отсутствуют работы по созданию комплексных методик и моделей анализа эффективности процедур тестового контроля. Проведение экспериментов на такой модели требует ее программной разработки с присущими ей методами планирования экспериментов по оценке как механизмов предъявления тестовых заданий, так и процедур классификации и оценивания.

Предметом исследования являются структура системы аттестации, включающая процессы обучения и методы компьютерного тестирования при организации сетевого режима обработки результатов аттестации сотрудников при процессах самотестирования, а также компоненты математического, лингвистического и программного обеспечения системы поддержки и принятия решений при классификации уровня знаний сотрудников транспортных предприятий.

1.2. Педагогические принципы контроля и диагностики

При обучении контроль и диагностика рассматривается как важнейшее, относительно самостоятельное и заключительное звено в управленческом цикле и является действенным механизмом выявления и оценки результатов проведенного действия. От правильной организации контроля и диагностики во многом зависит эффективность управления учебным процессом и качество подготовки специалистов. Обучение не может быть полноценным без объективной информации о том, как усваивается учебный материал, как обучаемые применяют полученные знания, умения и навыки для решения практических задач. В этих условиях следует рассмотреть сущность,

содержание, особенности и функции контроля в системе профессионального образования.

1.2.1. Функции и методы контроля

Контроль (от франц. *controle* – встречная, вторичная запись с целью проверки первой) – проверка, наблюдение; часто употребляется в смысле англ. *control* – господство, насилие, власть; это сравнение заданной, намеченной программы процесса обучения с фактическим ее выполнением, позволяющее:

- определять измерять и оценивать эффективность организации, методов и средств обучения, объем, глубину, осмысленность, осознанность и действенность приобретаемых обучаемыми знаний, умений и навыков;
- получать непрерывную информацию о ходе и качестве усвоения учебного материала, повышать эффективность познавательной деятельности обучаемых, выявлять ошибки и неверные действия и на этой основе проводить корректировку профессионального обучения;
- оценивать динамику усвоения учебного материала, действительный уровень овладения обучаемыми системой знаний, умений и навыков;
- создавать полноценные условия для развития личности обучаемого;
- стимулировать системную и целенаправленную работу обучаемых, активизировать их познавательную деятельность, определять уровень сформированных у них навыков обучения;
- формировать профессиональную направленность и мотивацию обучаемых;
- осуществлять регуляторное управление познавательной деятельностью обучаемых и вносить коррективы в процесс самообучения.

Основное назначение контроля состоит в формировании обратной связи, информирующей об адекватности функционирования системы ее конечным результатом [41]. Обратная информация о деятельности обучаемого в условиях самостоятельного обучения является практически единственным

оперативным средством управления познавательной деятельностью обучаемого, т.к. на ее основе изменяется отношение тьютора к процессу обучения, осуществляется коррекция индивидуальной образовательной траектории. Прямая и обратная связь – это механизм функционирования системы обучения.

Суть проверки результатов обучения состоит в выявлении, измерении и оценке уровня усвоения знаний обучающимся, который должен соответствовать образовательному стандарту. Контроль результатов профессионального обучения приобретает характер педагогической диагностики.

Традиционный контроль имеет глубокие исторические корни. В системе обучения общепринятыми считаются четыре основных этапа [1]:

Первый этап – предварительный контроль: имеет диагностический характер; служит необходимой предпосылкой для успешного планирования и руководства учебным процессом; позволяет определить начальный (исходный) уровень знаний, умений и навыков обучаемых, установить научную обоснованность целей обучения, проверить объективность требований, предъявляемых в ходе учебного процесса. На основании результатов предварительного контроля, проводимого в начале учебного цикла, вносятся коррективы в организацию профессиональной подготовки.

Второй этап – текущий контроль: один из основных видов проверки знаний, умений и навыков обучающихся. Ведущая задача текущего контроля – регуляторное управление познавательной деятельностью обучающихся и ее корректировка. Текущий контроль позволяет получать непрерывную информацию о ходе и качестве усвоения учебного материала и на этой основе оперативно вносить изменения в учебный процесс. Другими важными задачами текущего контроля являются стимуляция системной и целенаправленной работы обучающихся, определение уровня формирования у них навыков самообучения.

Третий этап – периодический (рубежный) контроль: позволяет определять качество изучения обучающимися конкретных разделов и тем учебных дисциплин. Такой контроль проводится несколько раз за срок обучения. Контроль охватывает значительные по объему разделы учебного курса. С помощью периодического (рубежного) контроля выявляются логические взаимосвязи учебных предметов. Периодический (рубежный) контроль включает функции сравнительного контроля. Сравнивая результаты обучения на его различных этапах, можно выявить и оценить динамику и эффективность профессиональной подготовки.

Четвертый этап – итоговый контроль: направлен на проверку конечных результатов обучения, выявление степени овладения обучающимися системой знаний, умений и навыков, полученных в процессе обучения; носит интегрирующий характер, позволяющий выявлять сформированные тенденции в процессе обучения, его эффективность и конкретные результаты. Итоговый контроль создает условия для глубокого анализа и коррекции системы обучения.

Организация контроля за учебным процессом связана с выбором и обоснованием оптимального сочетания разнообразных методов, форм, видов и способов контроля, с учетом особенностей технологии открытого обучения и конкретной учебно-педагогической ситуации.

Методы контроля – это способы диагностической деятельности, позволяющие осуществлять обратную связь в процессе обучения с целью получения данных о качестве обучения и эффективности учебного процесса. Они должны обеспечивать систематическое полное, достоверное и оперативное приобретение информации об учебном процессе с целью его коррекции.

Контроль качества и результативности учебного процесса имеет принципиальное значение, прежде всего, для самообучения, где контроль переходит в самоконтроль. Контроль является исходным основанием для

формирования самоконтроля как механизма саморегуляции человека во время общения [45].

В обучении самоконтроль определяется как сознательная самооценка и регулирование обучаемым своих действий в соответствии с поставленной целью и действующими правилами выполнения учебных заданий. Самоконтроль можно рассматривать как механизм обратной связи, обеспечивающий эффективность процесса управления собственной познавательной деятельностью.

Нельзя забывать еще об одной важнейшей функции самоконтроля – возможности выбора наиболее эффективных видов учебной деятельности в соответствии с результатами контроля. В целях сопоставления уровня развитости самоконтроля у обучаемых, необходимо установить его качественные и количественные показатели. Основными качественными критериями успешности формирования самоконтроля у обучаемых являются: время, глубина, полнота, объективность, надежность, правильность выполнения методов самоконтроля.

Организационные и содержательные особенности открытого дистанционного обучения, его новизна и актуальность потребовали формирования новой технологии контроля за результатами обучения.

Принципиальное отличие предлагаемой технологии от традиционной заключается в использовании интегральных контрольных заданий, их решение обеспечивает поэтапный контроль за познавательной деятельностью обучаемых на любом ее этапе. Использование новой технологии контроля дополняется самоконтролем обучаемых. Наиболее трудной проблемой в предлагаемой структуре новой технологии контроля является разработка критериев оценки результатов выполнения контрольных заданий (Четвертый этап).

Основное требование к разработке оценок – адекватность целям, средствам и содержанию контроля, в связи с чем вряд ли целесообразны единые критерии оценок. Для репродуктивного уровня контроля оценка

может быть выражена такими показателями, как полнота, обобщенность, системность, действенность и прочность знаний. Для продуктивного уровня контроля требуются другие критерии. Решение задач – это всегда продуктивная мыслительная деятельность, которая трудна для измерения. Наиболее доступный путь измерения и оценки – по проявлениям этой деятельности. Исходными критериями оценки продуктивного уровня контроля являются: правильность решения заданий, самостоятельность их выполнения, наличие развернутости доказательств, т.е. теоретическая обоснованность их решения.

Внедрение новой структуры контроля не означает отказ от традиционного контроля. В тоже время предлагаемая модель в значительно большей мере адекватна особенностям технологии открытого обучения, ориентированной на конечный результат.

Цели, принципы, функции, формы, методы, виды контроля связаны между собой, влияют друг на друга и выступают как целостная совокупность взаимосвязанных компонентов, т.е. как система, которая выполняет ряд важнейших функций: проверочную, обучающую, развивающую, методическую, диагностическую и воспитательную.

При проведении контроля целесообразно исходить из следующих принципов реализации контроля:

1. Полнота контроля, то есть максимальный охват обучаемых. Соблюдение этого принципа обеспечивает эффективную обратную связь.
2. Плановость и систематичность контроля с учетом его объема, содержания, уровня, затрат времени, средств и сроков проведения.
3. Открытость контроля. Отказ от скрытого контроля, который кроме обиды, досады и напряжения в отношениях с обучаемыми ничего существенного не приносит.
4. Целенаправленность контроля.
5. Позитивность контроля. Доведение его результатов до обучающихся в конструктивной и уважительной форме.

6. Достоверность результатов. Оценка сильных и слабых сторон личности обучаемого, интереса к учебе. Повышение личной ответственности за результаты обучения как обучаемого, так и тьютора.

7. Объективность контроля. Снижение субъективного влияния проверяющих на результаты контроля.

8. Дифференцированность контроля по уровню. Соответствие глубины (уровня) контроля поставленным целям.

9. Аналитичность контроля. Обеспечение глубины контроля с использованием методов научного анализа.

10. Диагностичность контроля. Выявление точно поставленного, аргументированного, оптимального диагноза негативных (позитивных) тенденций в развитии.

11. Системность контроля. Сосредоточение контроля на объектах, явлениях, факторах, выражающих общие закономерные тенденции.

12. Действенность контроля. Возможность использования результатов контроля для эффективной коррекции его объектов.

13. Автоматизация контроля, т.е. применение информационных средств с учетом целей контроля и специфики учебного материала.

14. Простота организационных форм проведения контроля. Контроль должен быть организован так, чтобы свести к минимуму затраты времени тьюторов и обучаемых, а также финансовые расходы на его проведение.

1.2.2. Классификация педагогических тестов

В мировой педагогической практике тестовые методы контроля уже давно используются как эффективные процедуры педагогической аттестации и контроля. Особую актуальность тестовые методы контроля приобретают при дистанционном обучении, когда применение традиционных методов аттестации становится либо не возможным, либо не эффективным в силу самих особенностей технологии дистанционного обучения.

Тестирование – один из наиболее действенных методов контроля качества открытого обучения. Его преимущество заключается в возможности проводить исследование на всех этапах учебного процесса, а также проводить его столько раз, сколько необходимо. Тестирование обеспечивает возможность получения объективного и достоверного "среза" состояния профессионального обучения, показателей, которые можно подвергнуть точному измерению, и на этой основе выявить истинный уровень качества профессиональной подготовленности специалистов [2].

Важным достоинством тестирования является ориентация на норму, что позволяет сопоставлять и сравнивать оценки, полученные при помощи теста. **Педагогический тест** - это система заданий специфической формы, позволяющая измерить уровень обученности учащихся, совокупность их представлений, знаний, и навыков на той или иной области содержания.

Педагогический тест рассматривается именно как система, как упорядоченное множество тестовых заданий. Задания - это те элементы из которых составляется педагогический тест.

Тест обычно состоит из задания и эталона, который является образцом последовательно и правильно выполненного задания.

Развитие теории и практики открытого обучения позволяет выделить и обосновать следующие основные функции тестирования:

- определение уровня умственного развития, интеллектуальных и специальных способностей обучаемых, степени их мотивации к учебной деятельности;
- выявление начального (исходного) уровня знаний, умений и навыков обучаемых;
- установление уровня развития психофизиологических свойств и качеств личности в профессиональной сфере; определение профессиональной пригодности; проведение профессиональной экспертизы и реабилитации;

- контроль качества усвоения обучаемыми учебного материала в процессе обучения и самообучения;
- установление уровней обучаемости, обученности, внимательности и работоспособности;
- изучение индивидуальных особенностей обучаемых;
- выявление умений обучающихся проводить аналитическую оценку процессов, явлений, фактов;
- определение соответствия содержания обучения и приобретаемых знаний, умений и навыков.

Анализ зарубежной и отечественной литературы показывает, что существуют несколько подходов к проблеме классификации тестов [36].

Во-первых, педагогические тесты классифицируются по методологии интерпретации результатов тестирования на нормативно-ориентированные и критериально-ориентированные.

Во-вторых, педагогические тесты различаются на гомогенные и гетерогенные.

В-третьих, педагогические тесты в системе профессионального образования подразделяются на виды: вступительные, текущие, на тесты промежуточной и итоговой аттестации.

В-четвертых, по форме предъявления педагогические тесты различаются на "бумажные" (бланковые), когда испытуемому предоставляется распечатка теста на бумаге, и на компьютерные, когда обучаемый должен отвечать на задания работая на компьютере.

Нормативно-ориентированный педагогический тест (по-английски norm-referenced test) позволяет сравнивать учебные достижения (уровень подготовки, уровень профессиональных знаний и умений) отдельных испытуемых друг с другом. Критериально-ориентированный педагогический тест (по-английски criterion-referenced test) позволяет оценивать, в какой степени испытуемые овладели необходимым учебным материалом.

Нормативно-ориентированные педагогические тесты используются для того, чтобы получить надежные и нормально распределенные баллы для сравнения тестируемых.

Критериально-ориентированные педагогические тесты применяются для того, чтобы интерпретировать результаты тестирования в соответствии уровнем обученности испытуемых на хорошо определенной области содержания.

Существенное значение приобретает использование критериально-ориентированных тестов в системе профессионального образования для текущей, промежуточной и итоговой аттестации учащихся. Преподаватели могут использовать результаты критериально-ориентированных тестов, чтобы правильно оценить уровень подготовки учащихся, а также определить качество тех или иных программ обучения. Тесты этого типа позволяют проводить мониторинг успеваемости учащихся, вовремя определять недостатки в подготовке учащихся. Преподаватели и руководители отмечают, что результаты критериально-ориентированных тестов особенно полезны при диагностике недостатков в обучении, а также в итоговой и промежуточной аттестации учащихся. Критериально-ориентированные тесты также используются для оценки различных образовательных программ.

Гомогенный педагогический тест основывается на содержании какой-либо одной дисциплины. При его разработке необходимо четко отслеживать, чтобы каждое задание не выходило по своему содержанию за рамки данной дисциплины. Гомогенный тест может быть как нормативно-ориентированным, так и критериально - ориентированным, в зависимости от цели его создания. Такой тип педагогического теста наиболее распространен в учебном процессе.

Гетерогенный педагогический тест основывается на содержании нескольких дисциплин. Гетерогенный педагогический тест является по своему существу междисциплинарным. В большинстве случаев каждое задание гетерогенного теста включает в себя элементы содержания

нескольких дисциплин. Одной из разновидностей гетерогенного теста является набор гомогенных тестов, т.е. гетерогенный тест может состоять из ряда гомогенных частей (субтестов). Наибольшее распространение гетерогенные тесты получили при итоговом контроле и при профессиональной аттестации кадров.

Виды тестов в системе профессионального образования. Как уже отмечалось, основными видами контроля над учебной деятельностью обучающихся являются предварительный, текущий, периодический (рубежный, поэтапный, промежуточный) и итоговый контроль. (Таблица 1.1.)

Предварительный тестовый контроль осуществляется перед началом учебного цикла и проводится по нескольким дисциплинам.

Таблица 1.1.

Виды тестового контроля в системе профессионального образования
России и их основные характеристики

Вид тестового контроля	Тип теста
Предварительный контроль	Нормативно-ориентированный, гомогенный
Текущий контроль	Критериально-ориентированный, гомогенный
Периодический (рубежный) контроль	Критериально-ориентированный, гомогенный
Итоговый контроль	Критериально-ориентированный, гетерогенный

Текущий тестовый контроль может осуществляться в ходе повседневной учебной работы. Этот вид контроля имеет большое значение для стимулирования у слушателей стремления к самостоятельной систематической работе и чувства ответственности за его результаты.

Периодический (рубежный, промежуточный) тестовый контроль проводится обычно после изучения логически законченной части (раздела,

модуля) программы. Он состоит в проверке учебной деятельности слушателей по освоению сравнительно большего объема материала. Поэтому к промежуточному педагогическому тесту предъявляются повышенные требования. Он должен обладать достаточно высокой надежностью и валидностью.

Коэффициент надежности теста – корреляционный коэффициент, который показывает, в какой мере совпадают результаты измерений, проведенных в одинаковых условиях. Понятие надежности непосредственно связано со стандартной измерительной ошибкой - информацией о том, между какими значениями полученной численной оценки находится истинное значение успеваемости индивидуума

Валидность измерения – характеристика показывающая то, что данная методика позволяет измерять действительно требуемые критерии (характеристики) исследуемого педагогического явления.

Валидность подразделяется на несколько типов:

1. содержательная валидность - экспертное подтверждение соответствия диагностического материала программе и основным целям обучения в контролируемой предметной области, согласованности результатов диагностики с другими независимыми формами контроля знаний;
2. критериальная валидность - достаточный уровень корреляции результатов тестирования по отдельным заданиям и по всему тесту в целом;
3. техническая валидность - обеспечение достаточного числа эквивалентных форм измерителей (вариантов заданий, вопросов), предотвращающих возможность механического заучивания правильных ответов.

Итоговый тестовый контроль осуществляется в конце каждого учебного года, а также при завершении изучения дисциплины (курса) с обязательным учетом результатов текущего и промежуточного (периодического) контроля. Итоговый тест совместно с другими аттестационными процедурами может существенно повысить качество подготовки специалистов

Наиболее перспективным для итоговой аттестации является использование гетерогенных междисциплинарных профессионально-ориентированных тестов, ориентированных на профессиональную деятельность специалиста.

1.3. Методики построения шкалы оценки уровня знаний

В работах [48, 57] развиты принципы, необходимые для индустриального подхода к образованию. Эти принципы основаны на использовании единой системы на основе унифицированной классификации предметного содержания образовательных курсов, и, как следствие – собственно обученности (совокупности знаний, умений и навыков обучающихся). При этом необходима:

1) унификация определяемого качества - обученности, шкалы обученности и структуры измерительных средств массового применения тестов обученности;

2) унификация дидактической классификации содержания обученности и, соответственно, структуры дидактических стандартов. Их однозначная связь с тестами обученности;

3) методологически унифицированная, но гибкая и оперативная аттестационная система — система установки образовательных планок (норм) органами управления образованием на основе объективных измерений качества образования с помощью массовых тестирований и конкретной общественной потребности (административные или социальные стандарты).

1.3.1. Принципы построения шкал в задачах тестового контроля

В настоящее время в образовании сосуществуют два основных вида оценки результатов учебной деятельности - субъективная, выставляемая преподавателем или группой (комиссией) экспертов, и объективная,

являющаяся результатом педагогического тестирования. Оба вида оценки необходимы и взаимодополняют друг друга.

Объективная тестовая оценка уровня обученности учащихся выражается в индивидуальном балле на определенной шкале.

Объективная тестовая оценка учебной деятельности учащихся имеет большое образовательное и воспитательное значение. Образовательная роль тестовой оценки состоит в том, что учащиеся получают объективную информацию о результатах своей учебной работы. Дополняемая самооценкой слушателя и субъективной оценкой преподавателя она указывает на достижения учащегося в овладении знаниями, умениями и навыками, а также на недостатки, пробелы, упущения и пути их устранения.

Количественным выражением субъективной оценки является отметка. Отметка условно выражает количественную оценку знаний, умений и навыков в цифрах или так называемых баллах. В системе ПО России применяется пятибалльная система отметок, а точнее четырех балльная система отметок. В этой системе субъективных отметок трудно обеспечить объективность и точность при оценке результатов учебной деятельности. Все разнообразие оттенков уровня подготовки невозможно вложить в жесткие рамки пяти баллов. Возможность расширить и более подробно дифференцировать диапазон оценок позволяет только применение тестовых технологий и внедрение в практику педагогических измерений тестовых шкал [32].

Достоверность педагогических оценок успешности обучения можно повысить, заменяя традиционные аттестационные процедуры стандартизованными педагогическими тестами с высокими показателями надежности и валидности.

В.С.Аванесовым [2] была предложена десятибалльная шкала, которая реализована в учебном пособии для институтов и факультетов повышения квалификации по основам педагогики и психологии высшей школы.

Для большей дифференциации и ухода от психологических трудностей преподавателей была предложена одиннадцатибальная шкала, которая получается из известной в литературе шкалы стандартных девяти единиц за счет добавления по одному интервалу в 0,5 стандартного отклонения справа и слева.

Соотношение между значениями нормированных отклонений Z , процентом испытуемых и оценочными эквивалентами можно представить в виде таблицы. (Таблица 1.2.)

Особенно четко в одиннадцатибальной шкале дифференцируется один процент лучших и худших. Если согласиться с тем, что один хороший специалист лучше ста плохих и средних, то ясно, что от этого одного процента зависит довольно много.

Таблица 1.2.

Соотношение между баллами, Z -значениями и процентом испытуемых.

Баллы	Оценочный эквивалент	Z - значение	% испытуемых
1	Низшая оценка	$-\infty < Z < -2,75$	1
2	Неудовлетворительно	$-2,75 < Z < -1,75$	3
3	Малоудовлетворительно	$-1,75 < Z < -1,25$	7
4	Удовлетворительно	$-1,25 < Z < -0,75$	12
5	Ниже среднего	$-0,75 < Z < -0,25$	17
6	Средняя оценка	$-0,25 < Z < 0,25$	20
7	Выше среднего	$0,25 < Z < 0,75$	17
8	Хорошо	$0,75 < Z < 1,25$	12
9	Очень хорошо	$1,25 < Z < 1,75$	7
10	Отлично	$1,75 < Z < 2,25$	3
11	Высшая оценка	$2,25 < Z < \infty$	1

Используя аттестационные критериально-ориентированные педагогические тесты, можно пытаться выйти за пределы порядковой шкалы и перейти в область интервальной шкалы.

Шкала, или шкалы, могут использоваться как набор норм, или как набор интерпретируемых данных, а для тестов, у которых есть несколько параллельных форм, как набор таблиц эквивалентности, которые позволяют преобразовывать необработанные баллы в общую шкалу для данного теста.

Шкалирование педагогического теста - это метод для определения системы преобразования индивидуальных необработанных баллов испытуемых в результате тестирования в баллы, которые окончательно сообщаются пользователям теста или тестовых программ.

Применение стандартизованных педагогических шкал по результатам тестирования позволит проводить статистическую обработку количественных показателей оценки учебной работы учащихся, что позволит с достаточной объективностью судить об успехах работы преподавателей и самих учебных заведений.

1.3.2. Формализованные модели классификации обученности

В структуре *знаний* для каждого образовательного уровня целесообразно выделить только четыре звена [45, 81]:

М - мировоззренческий минимум (это те знания, которые должны остаться в памяти любого обучающегося по данному предмету);

Б - базовые знания как дополнение к минимуму, которое необходимо для дальнейшего успешного изучения данного предмета;

П - программные знания сверх базового уровня;

С - сверхпрограммные знания, рекомендованные как дополнение к программе для самых сильных обучающихся.

В структуре *умений* также выделим четыре уровня:

Ф - фактический - предполагает умение узнавать (идентифицировать) основные факты, формулы, термины и принципы предмета;

О - операционный - предполагает выполнение действий по образцу (умение реализовать известный алгоритм);

А - аналитический - предполагает умение анализировать ситуацию и строить процедуры из простых освоенных операций;

Т - творческий - доступен будущему профессионалу, свободно владеющему материалом предмета и способному находить нетривиальные решения

Рассмотрим сначала структурирование знаний.

Когда обучающийся находится на том или ином уровне знаний (М, Б, П или С), он не просто "больше" или "меньше" знает однородных по своей логической структуре фактов (с этой точки зрения слова "интеграл" и "ромашка" логически однородны - это просто некие, хотя и совершенно различные по содержанию, понятия). Обучающийся обладает не просто тем или иным объемом информации (который можно выразить, например, в информационных единицах-битах), но эта информация обычно качественно различается по своей логической структуре, по уровню абстрагирования.

Поскольку мировоззренческий минимум выступает в роли обязательного элемента, требуемого образовательным стандартом в качестве необходимого минимума знаний любого человека, даже самые сложные объекты здесь выступают (подаются) как цельные и предельно простые сущности - предметы, качества или действия (операции).

Сверхпрограммные знания включают описание особенностей, важных для специалистов в данной предметной области, позволяют ввести и использовать межсистемные связи и системные комплексы. Общий объем информации на этом уровне существенно превышает уровень программы.

Рассмотрим экстенсивный компонент знаний. Он тиражирует знания четырех рассмотренных выше уровней - М, Б, П и С, наполняет их конкретным (предметным) содержанием. Поэтому очевидно, что знания, по структурной интенсивности соответствующие звену Б, не могут находиться в звене М, поскольку это сразу же изменит логический статус

мировоззренческого минимума, рассмотренного выше (М тем самым превратится в Б)

Совершенно аналогично знания со структурной интенсивностью П не могут входить в Б, а С - в П

Однако логико-структурный статус уровней знаний (звеньев) М, Б, П и С не нарушится, если на этих уровнях будут находиться знания, соответствующие им по содержанию, но логически структурированные на более низком уровне

Например, на уровне П могут находиться знания с логической структурой М или Б, содержание которых соответствует именно программному уровню (уровню хорошо обучающихся), поскольку на уровнях М и Б эти знания остались бы невостребованными. Мало того, знакомство с ними только запугало бы слабого обучающегося - затруднило бы освоение уровней М или Б (напомним, что на уровне М находятся знания, необходимые любому человеку)

Таким образом, экстенсивные знания любого из уровней Б, П и С могут включать (и, как правило, включают) знания структурно более низких уровней. Б включает знания с интенсивностью, соответствующей также и М. П - с интенсивностью (кроме П) еще Б и М. С — с интенсивностью (кроме С) М, Б, П

Смешение различных по логической структуре знаний на каждом уровне (кроме М, где находятся простейшие по структуре знания) создает определенные трудности при предметной классификации. Если понятно, что знания со сложной логической структурой не могут находиться на уровнях М или Б, то по какому принципу нужно размещать знания с низкой логической интенсивностью на высших уровнях, например, с интенсивностью типа М на уровнях П или С?

Базовый уровень обеспечивает минимум знаний, необходимых для дальнейшего освоения предмета в рамках программы

Сформулируем кратко функционально-логический статус выделенных

уровней знаний:

М - включает простые сущности, определяемые через их качества (свойства) и простейшие модели (схемы) данной предметной области, необходимые каждому человеку как гражданину;

Б - содержит основы знаний, позволяющие успешно овладевать предметом на текущей и последующих стадиях обучения;

П - включает спектр основных современных знаний, в том числе основные модели объектов и явлений данной предметной области.

С - содержит описания объектов и явлений данной предметной области, которые не входят в программу, но необходимы будущему специалисту в этой области.

Умения, как и знания, качественно различны. Во-первых, их можно разделить на две группы: репродуктивные (от учащегося преимущественно требуется воспроизведение показанного преподавателем) - таковы уровни Ф и О, и продуктивные, преимущественно требующие поиска самостоятельных решений, - таковы умения уровней А и Т. Во-вторых, умения каждого из четырех уровней (Ф, О, А, Т) в зависимости от знаний, на которых они базируются, отличаются в пределах каждого уровня по интенсивности совершенно аналогично тому, как мы это определили для знаний.

Из-за указанной выше логической связи знаний и умений их нужно рассматривать вместе. Для этого блоки знаний и умений удобно отобразить в виде таблицы (матрицы) на так называемой предметной плоскости или плоскости обученности (рис.1.2.).

Буквы указывают уровень обученности в данной предметной области, цифры характеризуют относительную сложность заданий.

Предметная плоскость может строиться как для традиционных предметов, так и для их комплексов, ориентированных на интеграцию отдельных предметов в области знаний (как это предусмотрено, например, базисными учебными планами общеобразовательной школы).

Плоскость (предметная плоскость) и матрица обученности

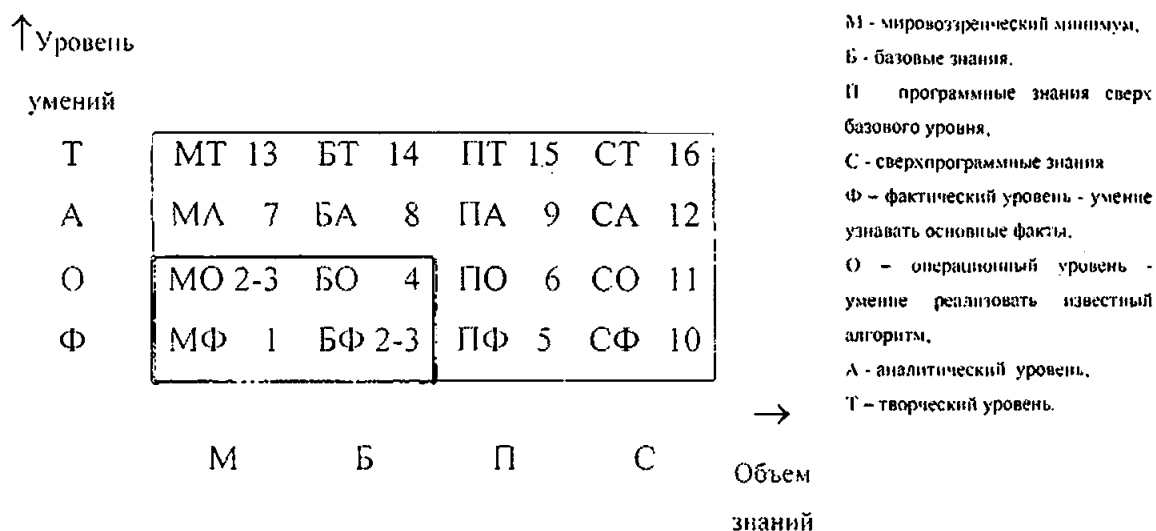


Рис. 1.2.

Элементы матрицы обученности на рис.1.2., обведенные в левом нижнем углу двойной линией (МО, МФ, БО, БФ), обычно относят к образовательному минимуму, определяемому образовательными стандартами. Здесь расположен материал, доступный слабым обучающимся.

Элементы матрицы обученности сверху и справа, отмеченные заполнением (МТ, БТ, ПТ, СТ, СА, СО и СФ), относятся к сверхпрограммному материалу (правый столбец) или к повышенному уровню трудности (верхняя строка с индексом Т). Здесь размещают материал для сильных обучающихся.

Ячейки без дополнительной обводки (МА, БА, ПА, ПО и ПФ) включают учебный материал предмета для хорошо подготовленных обучающихся.

Цифры в ячейках указывают на порядок нарастания сложности помещаемого в ячейках учебного материала.

Плоскость обученности наглядно демонстрирует неразрывную связь знаний и умений, их единство в более общем и важном для педагогической практики понятии обученности.

Рассмотрим последовательно все элементы матрицы на плоскости обученности.

МФ - самый простой по содержанию, но и самый общий, самый важный исходный элемент обученности. Сюда относится учебный материал об объектах, их качествах (свойствах) и о действиях с ними (операциях), которым учащиеся должны владеть на самом низком уровне - уровне узнавания, но владеть всегда! Например, будущий экономист может забыть всё, что он изучал по физике, кроме материала, помещенного в ячейке МФ.

Умение работать с объектами только мировоззренческого уровня по предписанным алгоритмам, повторяя преподанный образец, предусматривает учебный материал ячейки МО.

Знания о том, как совершать определенные программой простые действия с объектами мировоззренческого и базового уровней (законы, правила) и та часть новых объектов, которые не включены в мировоззренческий минимум, но необходимы (как и операции над ними) для дальнейшего обучения, входят в ячейку БФ. Учебный материал этой ячейки не предполагает умения данные действия совершать. О них нужно только знать.

Ячейка БО содержит учебный материал об объектах, качествах и действиях, знание которых и умение их использовать (по образцам) достаточно для дальнейшего изучения предмета. Обучающийся может воспроизводить алгоритмы преобразований для известных ситуаций.

Материал ячеек МФ, МО, БФ и БО идет по возрастанию его сложности под номерами 1-4.

В качестве обязательного элемента знаний ячейки ПФ присутствуют сведения о логике умозаключений, то есть правила проведения анализа фактов и сведения о том, как делают выводы по результатам анализа. Умение самому делать такие выводы и анализировать факты на этом уровне от обучающегося не требуется.

Содержание ячейки ПО предусматривает операционное владение элементами анализа (решение простейших логических задач по образцам).

Ячейки аналитического уровня МА и БА включают учебный материал для обучающихся, способных самостоятельно проводить анализ, ранее выполненный преподавателем по близким, но не тождественным образцам, делать вслед за ним выводы по материалам соответственно мировоззренческого и базового уровней.

При этом в ячейке МА размещается материал, по которому обучающийся должен уметь проводить доказательства и делать выводы (следствия) из фактов и операций мировоззренческого уровня, в ячейке БА - то же, что и в МА, но по материалу базового уровня. Кроме того, информация ячейки БА также предполагает, что преподаватель демонстрирует различные варианты выводов и доказательств, а обучающийся оценивает их преимущества (при возможной помощи преподавателя).

Ячейка ПА - то же, что и БА, но здесь содержится уже весь программный материал. Кроме того, предусматривается, что обучающийся проводит самостоятельный анализ новых для него сочетаний фактов и операций.

Соответствующая указанным ячейкам ПФ, ПО, МА, БА и ПА последовательность сложности материала на рис. 1 идет под цифрами 5-9.

Столбик из ячеек СФ, СО, СА заполняется информацией совершенно аналогично столбiku ПФ, ПО и ПА, но содержит сверхпрограммный материал для углубленного изучения предмета. Сверхпрограммный материал может быть как оригинальным, так и преподанным в других курсах (предметах), а в данном курсе устанавливается системная межпредметная связь.

Ячейка СФ содержит сверхпрограммный материал (факты и сведения об операциях) для углубленного изучения предмета. Этот материал может быть усвоен преимущественно обучающимися, овладевшими программными знаниями и аналитическими навыками, соответствующими ячейке ПА. Владения новыми операциями не требуется.

СО - то же, что в СФ, но требуется владение новыми операциями

Ячейка СА предусматривает исследование различных вариантов связей, доказательств, систематик поиск новых аналогий под руководством преподавателя

Строка Т" матрицы (ячейки МТ, БТ ПТ и СТ) отличается от строки "А" (ячейки МА БА ПА СА) только тем, что расположенный в ней материал предусматривает свободное владение знаниями соответствующих уровней и способность обучающегося делать самостоятельные выводы и прогнозы, проводить новые оригинальные доказательства, самостоятельно строить плодотворные аналогии, классифицировать материал и самостоятельно (ячейка СТ) приобретать новые знания

Иворческий уровень предусматривает умение обучающегося самостоятельно (без предварительного разъяснения преподавателя) проводить доказательства, анализ, выводы и следствия МТ - по материалу преимущественно мировоззренческого уровня. БТ — базового уровня. ПТ — по всей программе СТ - по всей программе с использованием новых приемов и сведений почерпнутых самостоятельно из других программ (других предметов) или из сверхпрограммного материала, изученного самостоятельно и по собственной инициативе обучающегося Эти ячейки соответствуют, как правило, возрастающей сложности содержащегося в них материала

Приведенная классификация обученности позволяет строить дидактически оптимизированные программы и контрольные процедуры к ним на основе единого, унифицированного подхода к учебному материалу

Унифицированная – в указанном выше ограниченном чисто квалификационном смысле – образовательная система приобретет не только способность контроля всех своих элементов но и возможность постоянного выявления и отбора из них для массового внедрения лучших наиболее совершенных и эффективных средств техники и технологий (методик) обучения, то есть собственных (образовательных) средств производства

1.3.3. Унифицированная дидактическая классификация

Одной из главных функций любого преподавателя является нормирование шкал обучающихся по собственным шкалам преподавателя. Если такого нормирования нет, обучающиеся просто не будут воспринимать требований преподавателя — они будут им непонятны, а ответы будут неадекватны контрольным вопросам.

Унифицированная дидактическая классификация (см. рис.1.2.) позволяет упорядочить учебный материал любого предмета с точки зрения оценивания обученности. Для этого в соответствии со сложностью материала можно развернуть в линию все ячейки данного рисунка, полученная последовательность ячеек, расположенная в порядке нарастания сложности материала (табл.1.3.):

Таблица 1.3.

Шкалы трудностей

МФ	МО	БФ	БО	ПФ	ПО	МА	БА	ПА	СФ	СО	СА	МТ	БТ	ПТ	СТ
1	2-3	2-3	4	5	6	7	8	9	10	11	12	13	14	15	16
10%	10%	10%	10%	10%	10%	10%	10%	10%	Трудность заданий						
Трудность заданий:					10%	20%	30%	40%	50%	60%	70%				

В третьей строке указана примерная трудность заданий данного типа для массового обучения (с точностью -10%). Для углубленного изучения предмета шкала трудности охватывает весь диапазон обученности (четвертая строка), кроме самого легкого материала.

Таким образом, плоскость обученности скомпонована в линию, и получена общая шкала для знаний-умений любого конкретного предмета.

Если при оценке знаний и умений с помощью любой формы контроля учитывается также время ответа обучающегося (а время на подготовку ответа практически всегда учитывается), контролируются не только знания и умения, но и навыки.

При специально вводимом ограничении времени подготовки ответа роль навыков возрастает с уменьшением этого времени и уменьшается с увеличением. Поэтому, в случае контроля времени подготовки ответа обучающегося, шкала для контроля знаний и умений превращается в шкалу обученности.

Предложено много способов численного выражения результатов измерения обученности. В России сегодня наиболее популярна пятибалльная шкала.

В практике нетестового (бесприборного) оценивания многобалльные - например, шестнадцатибалльные - шкалы нерациональны, поскольку оценивающий обычно не в состоянии осуществить разумный выбор одного из двух соседних баллов. Поэтому на шкале обученности целесообразно объединение интервалов для уменьшения числа используемых баллов.

Например, для применения традиционной пятибалльной шкалы оценок для массового обучения естественно следующим образом скомпоновать ячейки обученности (см. рис.1.2.):

1. "очень плохо" — обученность в пределах ячейки МФ;
2. "плохо" или "неудовлетворительно" - обученность за пределами ячейки МФ, но меньше минимума образовательного стандарта;
3. "удовлетворительно" - обученность соответствует ячейкам минимума образовательного стандарта МФ, БФ, МО и БО (номера сложности от 1 до 4 включительно);
4. "хорошо" - знания, умения и навыки охватывают материал предмета в ячейках МА, БА, ПФ и ПО (сложность от 5 до 8);
5. "отлично" - ячейки ПА, МТ, БТ, ПТ, СТ, СФ, СО и СА с номерами от 9 до 16 включительно.

Как показывает имеющийся педагогический опыт в нашей стране и за рубежом, пятибалльная шкала себя изживает. Например, сегодня одинаковые оценки вынуждено выставляются за совершенно разное качество обученности. Поэтому, целесообразно выставлять оценки по двум шкалам:

двухбалльной (зачет-незачет) и семибалльной (Международный бакалавриат) или одиннадцатибальной (по Аванесову В С)

Двухбалльная шкала хороша для обучающихся, которым в их будущей профессии не придется использовать знания соответствующего предмета. Освоившему минимум знаний, который определяет образовательный стандарт, можно ставить "зачет" (знания, умения и навыки выходят за пределы ячеек МФ, МО, БФ и БО на рис 1 2)

Семибалльная шкала позволяет сохранить и использовать традиционные пять баллов для обычных общеобразовательных программ, а шестой и седьмой - для программы углубленного изучения предмета. Такая шкала позволяет объективизировать оценку по результатам тестирования и отказаться от вынужденно применяемой "гибкой" методики оценивания.

Например, можно реализовать следующий вариант градуировки

- 1 "очень плохо" - знания не выходят за пределы ячейки МФ.
- 2 "плохо" или "неудовлетворительно" - знания превышают МФ, но не выходят за пределы БФ.
- 3 "Удовлетворительно" - обученность в пределах стандарта (совокупность ячеек МФ, БФ, МО и БО).
- 4 "хорошо" - обученность вышла за пределы минимума образовательного стандарта и охватывает ячейки ПФ и ПО.
- 5 "отлично" - обученность полностью соответствует общеобразовательной программе (ячейки МФ, БФ, ПФ, МО, БО, ПО, МА, БА, ПА).
- 6 "успешно" - обученность "сверхпрограммная" - соответствует программе углубленного изучения предмета (включаются 9 ячеек, указанных выше, плюс ячейки СФ, СО, СА).
- 7 "превосходно" - демонстрируется обученность по всем 16 клеткам матрицы

1.4. Математическое моделирование процедур тестового контроля

Под тестированием понимается процесс организации такого целенаправленного воздействия на субъект, в результате которого осуществляется оценка знаний, умений и навыков обучаемого. Основными подходами к построению математических моделей тестового контроля можно считать: классическую теорию; IRT-теорию; факторный анализ; латентно структурный анализ; кластерный анализ.

1.4.1. Основные принципы классической теории тестового контроля

Классическая теория основана на обработке статистических данных по результатам тестовых результатов. В качестве модели применяет простейшая аддитивная модель, в которой на истинный бал накладывается ошибка. В классической теории тестов зарубежные и российские авторы исходили из идеи параллельного измерения интересующего свойства с помощью двух и большего числа тестов, имеющих общее предметное содержание и сходные статистические характеристики [78]. В рамках этой теории приняты следующие постулаты:

Эмпирически получаемый результат измерения (X) представляет собой сумму истинного компонента измерения (T), обычно неизвестного, и ошибочного (E), также неизвестного:

$$X=T+E. \quad (1.1)$$

Истинный компонент измерения можно выразить как математическое ожидание $E(X)$. Ошибочный компонент измерения представляет собой результат влияния случайных ошибок. Среднее арифметическое этих ошибок принимается равным нулю. Дисперсии ошибочных компонентов параллельных тестов равны. Корреляция истинных и ошибочных компонентов по множеству испытуемых равна нулю.

Трудность задания определяется эмпирически и в последнее время она соответствует доле неправильных ответов:

$$q_i = \frac{W_i}{N}, \quad (1.2)$$

W_i - число неправильных ответов

N - число испытуемых

В теории тестов и в практике тестирования традиционно рассматриваются два основных критерия, по которым оценивается качество теста. Первый – это надежность тестов, ассоциируемая обычно с идеей точности измерения. Второй критерий – это валидность теста, определяемая обычно как способность теста измерять то, что он призван, по замыслу авторов, измерить. При измерении надежности предполагается линейная модель, в которой дисперсия балла равна сумме истинного и ошибочного.

Формула Kuder-Rishardson (KR-20) определяет коэффициент надежности при однократном тестировании:

$$r = \frac{K}{K-1} \left(1 - \frac{\sum p_i q_i}{S_A^2} \right), \quad (1.3)$$

где $\sum p_i q_i$ - сумма дисперсий заданий теста;

S_A^2 - дисперсия тестовых баллов испытуемых;

K – число заданий в тесте.

Формула D.Armor определяет коэффициент надежности теста как:

$$\theta = \frac{K}{K-1} \left(1 - \frac{1}{\lambda_1} \right), \quad (1.4)$$

λ_1 - наибольшее значение корня характеристического уравнения $|R - \lambda I| = 0$

Эффективным называется тест, который лучше других измеряет знания интересующего уровня подготовленности меньшим числом заданий, качественнее, быстрее, дешевле, и все это, по возможности, одновременно; эффективный тест представляет собой не одну какую-либо фиксированную систему заданий, а множество систем заданий, каждая из которых является оптимальной для измерения знаний на требуемом уровне подготовленности.

1.4.2. Формальные модели IRT-теории тестового контроля

Более формальной является Item Response Theory (IRT) [63], в которой предполагается существование аналитических выражений (в общем случае нелинейных) для задания вероятностей правильных ответов тестируемых определенного уровня за задачи определенного уровня сложности. Такой подход при описании модели тестируемого позволяет использовать эту модель во множестве аспектов. для модели регрессионного анализа, где численными методами оценивается неизвестный латентный параметр; для моделей классификации, так как ITR-модель дает возможность вычислить параметры распределений гипотетических групп классификации; для использования в имитационных моделях для генерации результатов ответов тестируемого (метод Монте-Карло).

IRT нацелена на оценивание латентных качеств личности и параметров заданий теста на основе математико-статистических моделей. К наиболее значимым преимуществам IRT обычно относят: устойчивые объективные оценки параметра, характеризующие уровень знаний испытуемых; устойчивые, объективные оценки параметра трудности заданий, не зависящие от свойств выборки испытуемых, выполняющих тест; измерение значений параметров испытуемых и заданий теста в одной и той же шкале, имеющей свойства интервальной шкалы, возможность предсказания вероятности правильного выполнения теста любым испытуемым в выборке до предъявления теста группе, возможность оценить эффективность различных по трудности заданий для измерения данного значения латентного параметра испытуемых

В IRT рассматривается условная вероятность правильного выполнения i -ым испытуемым с уровнем знаний Θ_i различных по трудности заданий теста, считая Θ_i параметром i -го обучаемого, а β - независимой переменной. В этом случае условная вероятность будет функцией латентной переменной β :

$$P_i\{x_{ij}=1|\Theta_i\}=f(\Theta_i - \beta) \quad i=1..N \quad (1.5)$$

Аналогично вводится условная вероятность правильного выполнения j -го задания, трудности β_j различными испытуемыми группы. Здесь независимой переменной является Θ , а β_j – параметр, определяющий трудность j -го задания теста:

$$P_i\{x_{ij}=1|\beta_i\}=f(\Theta - \beta_i) \quad i=1..n. \quad (1.6)$$

Число параметров, входящих в аналитическое задание функций, является основанием для подразделения семейства IRT на классы. Среди логистических функций различают однопараметрическая модель G.Rasch, двухпараметрическую модель A.Birnbaum, которые являются частными случаями трехпараметрической модели A.Birnbaum:

$$P(\Theta, \beta, a, c) = c + (1 - c) \frac{\exp(1,7a(\Theta - \beta))}{1 + \exp(1,7a(\Theta - \beta))}. \quad (1.7)$$

В каждой из представленных моделей параметры Θ и β выражаются как шкалированные показатели единой для всех моделей шкалы логитов. Введение единой шкалы для элементов двух различных множеств: значений Θ и значения β — позволяет решить ряд вопросов как теоретического, так и практического характера: появляется возможность ввести взаимосвязь между переменными в виде разности $|\Theta - \beta|$, корректно сравнить результаты, полученные с помощью различных тестов, оценить трудность заданий теста независимо от уровня подготовленности групп испытуемых. Эти важные преимущества позволяют преодолеть ряд отмеченных выше существенных недостатков классической теории тестов.

По сравнению с классической теории тестирования в IRT теории используется другая мера трудности:

$$\beta_j = \ln \frac{q_j}{p_j} \text{ - мера трудности задания (IRT) или логит трудность задания.}$$

Уровень знаний испытуемого в IRT теории определяется как:

$$\theta_i = \ln \frac{p_i}{q_i} \text{ - уровень знания } i\text{-го испытуемого.}$$

1.4.3. Факторный, латентно-структурный и кластерный анализ в системе моделирования процедур тестового контроля

Факторный анализ, так же как и IRT-теория строит модель тестируемого и позволяет моделировать его ответы в зависимости от сочетания параметров тестируемого и сложности задания. В этом подходе модель является линейной по параметрам. Проблемы возникают в объяснении значимости гипотетических (латентных факторов). Основной постулат факторного анализа выражается в виде простого линейного уравнения:

$$Z_{ij} = a_{j1}Z_{i1} + a_{j2}Z_{i2} + \dots + a_{jq}Z_{iq} \quad (1.8)$$

где Z_{ij} - есть стандартный балл индивида i по тесту j ;

$Z_{i1}, Z_{i2}, \dots, Z_{iq}$ - стандартные баллы индивида по гипотетическому набору статистически независимых характеристик или факторов.

Коэффициенты a являются набором весов для теста j и инвариантны по отношению к индивидам.

Если факторный анализ применяется к серии тестов, существенно различающихся по трудности, но совершенно очевидно зависящих от одной латентной величины, то возникает не один, а несколько факторов, по одному на каждую ступень трудности. Это в основном приписывается нелинейным соотношениям между тестами, существенно различными по трудности, причем эта нелинейность вызвана дифференциальной асимметрией распределений баллов. Отметим, однако, что факторная модель накладывает ограничение линейности только на соотношения между опытами и факторами, как это видно из формулы (1.8).

Процесс изменения "уровня знаний" учащегося можно рассматривать как многомерный случайный процесс его "баллов" по различным направлениям специализации. И в общем случае такая модель необходима при проверке робастности статистических процедур.

В общем случае факторная модель представляет:

$$\xi_i = \sum_{j=1}^m \lambda_{ij} F_j + \varepsilon_i, \dots, \xi_p = \sum_{j=1}^m \lambda_{pj} F_j + \varepsilon_p, \quad (1.9)$$

где λ_{ij} - факторные нагрузки $i, j=1..m, m < p$;

$F_1, F_2, \dots, F_p$ - общие факторы, причем $DF_i=1$; $\text{cov}(F_i, F_j)=0 \quad i \neq j$

$\varepsilon_1, \dots, \varepsilon_p$ - специфические факторы, причем $\text{cov}(\varepsilon_i, \varepsilon_j)=0 \quad i \neq j$

$D\varepsilon_i=\tau_i, i=1..p, \tau_i$ - специфическая дисперсия;

Кроме того, независимыми полагаются общие и специфические факторы. Соотношения для дисперсий определяют новую структуру или "факторизацию" дисперсий и ковариаций исходных переменных. В результате, дисперсия и ковариация представляются в виде функций от α_j и дисперсий главных компонент.

Латентно-структурный анализ в отличие от предыдущих не опирается на построение модели тестируемого и вероятности правильных ответов. Его основная сущность состоит в группировке тестируемых в однородные группы по отношению к результатам ответов на каждое тестовое задание. Суть однородности заключается в существовании центра группы и равновероятности ответов каждым тестируемым на каждое задание. Интерпретация комплектации групп ложится полностью на исследователя. Кроме того, в латентно-структурном анализе существует проблема неоднозначности решения, что приводит к дополнительным трудностям формирования однородности групп.

Латентно-структурный анализ является методом исследования внутренних соотношений между дихотомическими признаками, такими, как ответы на вопросы анкеты [61]. Метод основан на составлении линейных уравнений рескрутирования следующего вида:

$$\begin{aligned} n &= n_1 + n_2 + \dots + n_q \\ n_i &= n_1 \cdot p_{1i} + n_2 \cdot p_{2i} + \dots + n_q \cdot p_{qi} \\ n_{ik} &= n_1 \cdot p_{1ik} + n_2 \cdot p_{2ik} + \dots + n_q \cdot p_{qik} \\ n_{k1} &= n_1 \cdot p_{1k1} + n_2 \cdot p_{2k1} + \dots + n_q \cdot p_{qk1} \end{aligned} \quad (1.10)$$

Величины, находящиеся слева, являются эмпирически данными, или наблюдаемыми. Они соответствуют числу людей во всей выборке n , числу людей, ответивших положительно на один вопрос n_j , числу людей, ответивших положительно на два вопроса n_{jk} , и т. д. Величины справа - подлежащие дальнейшему рассмотрению латентные параметры модели. Количество величин справа q равно числу подгрупп (латентных классов), на которые в результате анализа будет производиться разделение всей выборки. Число людей в латентном классе l равно n_l т. д. Латентная вероятность p_{lj} равна доле членов латентного класса l , давших положительный ответ на вопрос j , вероятность p_{ljk} равна доле членов латентного класса l , давших положительные ответы на вопросы j и k , и т. д. Уравнения (1.10) просто показывают, каким образом эмпирические данные рекрутируются из латентных классов. По ходу дела можно заметить, что уравнения (1.10), которые являются системой уравнений рекрутирования, линейны по своей природе, в то время как исходные уравнения факторного анализа линейны просто потому, что их такими выбрали. Различие между наблюдаемыми и латентными характеристиками, которые так выпукло обнаруживаются в теории латентной структуры, являются, конечно, центральным и в факторном анализе. Тесты и их взаимные корреляции являются эмпирическими, а факторы и их нагрузки - латентными величинами. То же отличие имеет место и в модели латентного профиля.

Естественно требовать, чтобы каждый латентный класс был однородным относительно любых исследуемых (латентных) величин, могущих влиять на наблюдаемые соотношения. Полная однородность не обязательна, если отклонения от среднего в классе случайны.

Кластерный анализ, также как и латентно-структурный не строит модели отношения тестируемого к заданию определенной сложности. Основная задача – разбиение на группы. Однако в этом случае для реализации механизма группировки вводятся меры различия и меры сходства. Целью является, при заданном количестве групп, заранее

определенных исследователем, собрать в каждую наиболее близкие по введенной мере сходства, а в различных группах должны находиться тестируемые имеющие наибольшие меры различия. Реализация метода осуществляется итерационной процедурой, поэтому ей присущи все проблемы поисковых и итерационных алгоритмов.

Задача кластерного анализа заключается в разбиении множества объектов I на m кластеров (подмножеств) $I_1, I_2, \dots, I_m$ так, чтобы каждый объект принадлежал одному и только одному подмножеству разбиения. Цель разбиения - объединить сходные объекты в одном кластере и разнести в разные кластеры различающиеся объекты.

Решением задачи кластерного анализа является разбиение, удовлетворяющее некоторому критерию оптимальности. Этот критерий может представлять собой некоторый функционал, выражающий уровни желательности разбиений и группировок.

В отличие от латентно-структурного анализа в кластерном анализе вводится численная мера различия (через расстояние). Расстояние $d(X, Y)$ между множествами X и Y определяется величиной:

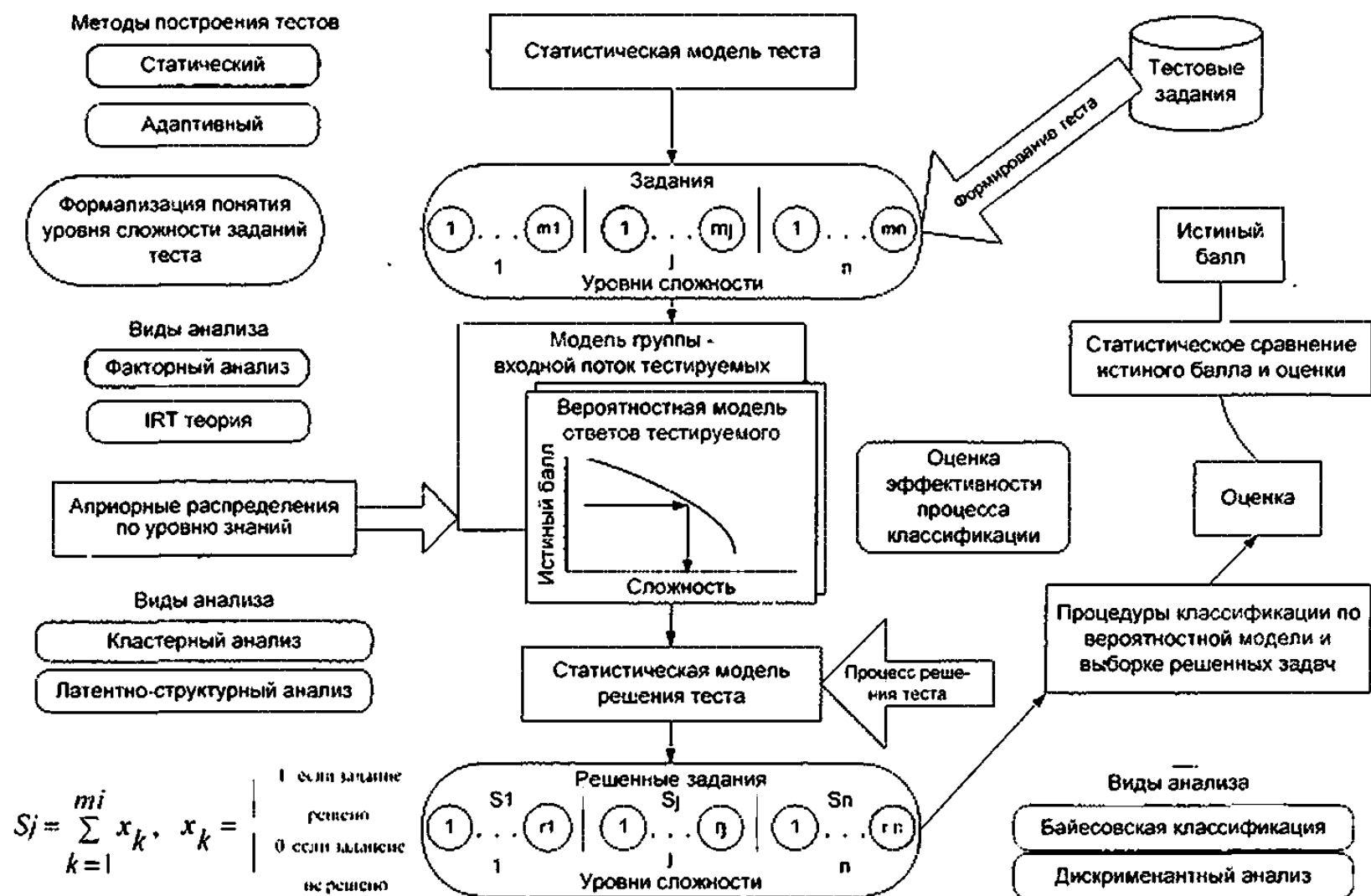
$$d_2^2(x, y) = \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} d^2(x_i, y_j). \quad (1.11)$$

Для оценки плотности расположения точек внутри множества используется мера близости. Общая идея, лежащая в основе подхода, использующего понятие близости, состоит в преобразовании пространства описаний D в пространство D^* , в котором все точки одного множества расположены близко друг к другу, а точки различных множеств удалены друг от друга на некоторое расстояние.

Для решения задачи кластерного анализа необходимо количественно определить понятие сходства и различия. Решением задачи является разбиение, в котором расстояние между X_i и X_j в одном кластере достаточно мало, а расстояние между X_i и X_j в разных кластерах достаточно велико.

Рис. 13.

Взаимосвязь методов и моделей тестового контроля



Основные усилия в развитии методов кластеризации и классификации были направлены на построение методов, основанных на минимизации внутригрупповых сумм квадратов отклонений. Они могут быть выражены в терминах евклидовых расстояний и называются методами минимальной дисперсии

В общем случае оценивание уровня знаний по результатам ответов на тестовые задания в диссертации рассматривается как задача управления, где на каждом шаге процедуры выбирается очередное задание, обладающее определенной сложностью. При формировании процедуры предъявления заданий варьируется выбор сложности по результатам ответов:

$$\beta^{(n+1)} = F^{(n)}(\beta^{(1)}, \dots, \beta^{(n)}) + \xi^{(n)}(\beta^{(1)}, \dots, \beta^{(n)}), \quad (1.12)$$

где $\beta^{(n)}$ - сложность задания на n -ом шаге процедуры;

$F^{(n)}$ - некоторое функциональное преобразование результатов ответов.

$\xi^{(n)}(\beta^{(1)}, \dots, \beta^{(n)})$ - случайная величина, моделирующая ответ на n -е задание.

В результате проведенного анализа показано, что при обучении контроль рассматривается как важнейшее, относительно самостоятельное и заключительное звено в системе переподготовки кадров.

Выводы по главе I

1 Рассмотрены проблемы кадрового обеспечения предприятий промышленности и транспортного комплекса, непрерывности образования и основные тенденции развития системы переподготовки и аттестации

2 Выделены основные педагогические принципы формирования технологии формирования тестов и тестовых заданий в системе аттестации. Рассмотрены педагогические принципы построения процедур контроля и диагностики. Приведены подходы к организации методической работы в учебных центрах

3. Проведена классификация существующих математических методов моделирования процесса тестового контроля и организации процедур аттестации и тестового контроля.

4. Рассмотрены информационные технологии, используемые при организации аттестации и тестового контроля в системе подготовки и переподготовки кадров.

2. ПОСТРОЕНИЕ МОДЕЛЕЙ КЛАССИФИКАЦИОННЫХ ХАРАКТЕРИСТИК И ОЦЕНОЧНЫХ ШКАЛ СИСТЕМЫ ТЕСТОВОГО КОНТРОЛЯ

Во второй главе диссертации рассматриваются вопросы построения функциональных соотношений описания индивидуальных свойств обучасмого, которые базируются на моделях восприятия и забывания учебной информации. С целью осуществления мониторинга уровня сформированности требуемых характеристик специалиста предлагается использовать комплексный показатель качества психолого-индивидуальных компетенций

$Kq = \sum_{i=1}^b Ka_i / b$, где Ka – локальные коэффициенты сформированности профессиональных компетенций, b – число локальных коэффициентов.

Психолого-индивидуальные компетенции представляют собой способности, развиваемые в профессиональной деятельности под влиянием мотивации, которая может, как усиливать, так и ослаблять потенциальные задатки специалиста. В этой связи необходимо осуществлять постоянный мониторинг мотивационной направленности. Для определения критериев сформированности проводится тестирование выборки специалистов, успешных в своей деятельности. Для визуализации локальных показателей используется их представление в виде профилограммы, где приняты обозначения: A – коэффициент точности внимания; E – коэффициент продуктивности влияния; $K_{расн}$ – коэффициент распределения внимания; V_m – коэффициент объема памяти; K_m – коэффициент творческого мышления; K_n – коэффициент логического мышления; $K_{л}$. Применение такой модели специалиста на основе процессного подхода и методологии «развертывания функции качества» позволяет снизить уровень неопределенности идентификации и мониторинга рассматриваемых компетенций и повысить качество переподготовки специалистов.

2.1. Модель дифференцированной оценки сложности тестового задания

В основном в системах тестового контроля считается два варианта интерпретации ответа на задание – решено, не решено. В работе предлагается методика формирования количественной оценки корректности решения тестового задания.

Пусть V – множество компонентов ответа, $V \neq \emptyset$, V_s – подмножество компонентов, выбранных обучаемым, $V_s \subseteq V$; V_r – подмножество компонентов, соответствующих верному результату $V_r \subseteq V$; V_a – подмножество компонентов, бессмысленных в контексте УТЗ. Подмножества V_r и V_a будут выделяться при конструировании УТЗ. В такой постановке правила оценивания определяют следующие варианты интерпретации:

- если $V_s = V_r$, то ответ верный;
- если $V_s \subseteq V_r$ & $V_s \neq \emptyset$, то ответ неточный или неполный;
- если $(V_s \setminus V_r \neq \emptyset \text{ \& } V_s \setminus V_a = \emptyset) \vee (V_s = \emptyset \text{ \& } V_r \neq \emptyset)$, то ответ ошибочный;
- если $V_s \cap V_a \neq \emptyset$, то ответ абсурдный.

Наборы оценок для разных УТЗ могут отличаться по полноте. Обязательными их элементами являются оценки верного и ошибочного результатов.

Дифференцирование решения УТЗ по сложности позволит увеличить дискретизацию этой шкалы, что будет выражено в интегральной рейтинговой оценке.

Так, естественное расширение закрытой формы – *множественный выбор*. Пусть: $A = \{A_m\}_{m=1..M}$ – множество альтернативных вариантов ответов; $ZA = \{ZA_m\}_{m=1..M}$, $ZA_m \in \{0, 1\}$ – вектор правильного выбора (1 – альтернатива входит в выбор, 0 – нет); $TA = \{TA_m\}_{m=1..M}$, $TA_m \in \{0, 1\}$ – вектор выбора тестируемого (1 – альтернатива выбрана, 0 – нет). В качестве расхождения предла-

гается использовать обычную равномерную метрику, но нормированную по количеству альтернатив:

$$\rho(ZA, TA) = 1 - \frac{1}{M} \sum_{i=1}^M |ZA_i - TA_i|. \quad (2.1)$$

Для заданий на *соответствие* и установление правильной последовательности (по оценке правильности они идентичны) введена другая мера. Левая часть задания определяет первоначальную перестановку (без ограничения общности ее можно считать возрастающей) $A = \{A_m\}_{m=1}^M$ – также множество альтернативных вариантов ответов; $ZA = \{ZA_m\}_{m=1}^M$, $ZA_m \in \{1 \dots M\}$ – правильная перестановка; $TA = \{TA_m\}_{m=1}^M$, $TA_m \in \{1 \dots M\}$ – перестановка тестируемого. В данном случае процент правильности предлагается определять на основании:

$$\rho(ZA, TA) = 1 - \frac{1}{M} \sum_{i=1}^M < ZA_i = TA_i > \quad (2.2)$$

Логическое условие $< \dots = \dots >$ принимает значение 1 при равенстве аргументов и 0 в противном случае.

Соотношения (2.1) и (2.2) дает значение 1 при полном соответствии ответа, 0 – при полном расхождении. Промежуточные значения определяют процент правильности решения тестового задания. Аналогично задаются меры различия для заданий на кластеризацию объектов и др.

Абсурдный ответ	Ошибка	Неполный ответ	Верный ответ
0			1

Рис. 2.1. Дифференциация оценочной шкалы ответа на УТЗ

Такая количественная оценка может быть использована для формирования интегральной оценки по всему тесту (рис.2.1.) При разделении шкалы [0,

1] на четыре области, каждому ответу можно приписать одну из четырех оценок: верный - 1, неточный - 0,8, ошибочный 0,5 и абсурдный - 0.

Такой подход к оценке результатов ответа на УТЗ предлагается использовать и для расширения модели адаптивного тестового контроля, где сложность следующего задания выбирается на основании решения предыдущего. Так, неточный ответ на УТЗ определенной сложности можно рассматривать как решение УТЗ пониженной сложности.

Расчет процента правильности решения с учетом дифференцированной шкалы ответа на УТЗ предлагается вычислять

$$P(\Theta | R, \beta, a, c) = c + (1 - c) \frac{\exp(1,7a(\Theta - \beta \cdot R))}{1 + \exp(1,7a(\Theta - \beta \cdot R))}, \quad (2.3)$$

где R определяет балл в шкале $[0, 1]$ по ответу на тестовое задание, что по IRT-модели просто снижает уровень сложности задания.

2.2. Параметризация логистических моделей Раша и Бирibaума

Простейшая модель функции успеха имеет вид

$$p = \frac{s}{s + t} \quad (2.4)$$

где s - уровень подготовленности участника тестирования, t - трудность тестового задания, а p - вероятность правильного выполнения задания. На практике аргументы $s \in (0; \infty)$ и $t \in (0; \infty)$ часто удобно выражать в логарифмическом масштабе. Введем для этого следующие обозначения

$$\ln s = \Theta, \quad \ln t = \delta \Leftrightarrow s = e^{\Theta}, \quad t = e^{\delta} \quad (2.5)$$

При этом функция успеха (2.4) принимает вид

$$p = \frac{e^{\Theta}}{e^{\Theta} + e^{\delta}} = \frac{1}{1 + e^{\delta - \Theta}} = \frac{1}{1 + \exp[-(\Theta - \delta)]} \quad (2.6)$$

и называется основной логистической моделью Раша. Здесь p вероятность того, что участник тестирования с уровнем подготовленности Θ правильно выполнит задание трудности δ . Очевидно, что аргументы $\Theta \in (-\infty; \infty)$ и $\delta \in (-\infty; \infty)$ измеряются одной и той же шкалой с единицей измерения один

логит. Вероятность успеха, как видим, зависит, по существу, только от одного параметра - разности $\Theta - \delta$, и потому модель (2.6) является однопараметрической. При этом

$$\lim_{(\Theta - \delta) \rightarrow +\infty} p = 1, \quad \lim_{(\Theta - \delta) \rightarrow -\infty} p = 0 \quad (2.7)$$

$$p = 0,5, \quad \text{если} \quad \Theta = \delta \quad (2.8)$$

Вероятность p как функция δ при фиксированном значении уровня подготовленности $\Theta = \Theta_0$ полностью описывает потенциальные возможности индивидуума с уровнем подготовленности Θ_0 при выполнении заданий всевозможных трудностей δ и потому представляется характеристической функцией уровня подготовленности Θ_0 . Вероятность p как функция Θ при фиксированном значении $\delta = \delta_0$ с той же полнотой характеризует возможности участников тестирования с различным уровнем подготовленности при решении задания трудности δ_0 и называется характеристической функцией трудности δ_0 .

Характеристическая кривая уровня подготовленности представляет собой стандартную характеристическую кривую уровня подготовленности 0, сдвинутую (без деформации) вдоль оси абсцисс на Θ_0 единиц. Это означает увеличение (если $\Theta_0 > 0$) или уменьшение (если $\Theta_0 < 0$) вероятностей успешного выполнения индивидуумом с $\Theta = \Theta_0$ заданий всевозможных трудностей. Следовательно, характеристические кривые, соответствующие различным уровням подготовленности, не пересекаются. Это выгодно отличает функцию (2.6) от других функций успеха, где пересечение возможно.

Аналогичная ситуация имеет место и при параллельном переносе характеристической кривой трудности δ_0 .

Функция успеха (2.6) принадлежит известному в теории вероятностей семейству логистических функций распределения вероятностей вида [23]

$$F(x) = \left(1 + e^{-ax}\right)^{-1} \quad (2.9)$$

с дисперсией $D(x) = \frac{1}{3} \left(\frac{\pi}{d} \right)^2$, где d - параметр. Точка перегиба всех соответствующих кривых имеет координаты $(0; 1/2)$, но крутизна кривых в точке перегиба пропорциональна величине параметра d . Этот факт делает целесообразным заменить однопараметрическую функцию успеха (2.6) двухпараметрической моделью вида

$$p = \left[1 + \exp(-d(\Theta - \delta))^{-1} \right] \quad (2.10)$$

В рамках этой модели каждому тестовому заданию определенной трудности $\delta = \delta_0$ соответствует не единственная характеристическая кривая, а семейство кривых, пересекающихся в единой точке перегиба $(\Theta = \delta_0; 1/2)$. В этой точке крутизна каждой кривой семейства, т.е. величина тангенса угла наклона касательной, равна $d/4$. Поэтому значение параметра d является дополнительной характеристикой задания. Большее значение d является предпочтительным. При малом d характеристическая кривая трудности δ_0 является пологой. На практике это означает, что участники тестирования с хорошим и плохим уровнем и подготовки выполняют данное задание с приблизительно равным успехом. Наоборот, если d велико, то шансы успешного выполнения данного задания участниками с $\Theta < \delta_0$ и $\Theta > \delta_0$ отличаются существенно. Поэтому параметр d принято называть дифференцирующей или дискриминационной способностью тестового задания.

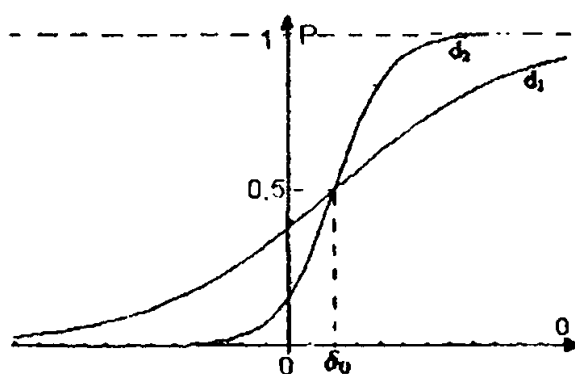


Рис. 2.2. Характеристические кривые трудности задания δ_0 с разными значениями параметра d ; $0 < d_1 < d_2$

Двухпараметрическая модель (2.10) (латентные параметры Θ , δ_0 и d) носит имя Бирнбаума [36]. При $d=1$ логистическая модель Бирнбаума (2.10) совпадает с логистической однопараметрической моделью Раша (2.6). При $d=1,7$ функция (2.10) хорошо аппроксимирует функцию распределения вероятностей нормального закона с нулевым математическим ожиданием и единичной дисперсией. А именно, если

$$N(\Theta / \delta = \delta_0; d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d(\Theta - \delta_0)} e^{-x^2/2} dx, \quad (2.11)$$

то

$$|N(\Theta / \delta = \delta_0; d) - p(\Theta / \delta = \delta_0; d)| < 0,01 \quad \text{для} \quad \forall \Theta \in (-\infty; \infty) \quad (2.12)$$

где $p(\Theta / \delta = \delta_0; d)$ обозначает характеристическую функцию трудности δ_0 в модели Бирнбаума с $d=1,7$. Этот факт позволяет различные вычисления с логистической моделью интерпретировать, при необходимости, с позиции детально изученного нормального закона распределения вероятностей.

Известна и трехпараметрическая модель Бирнбаума [36]. Она имеет вид:

$$p(\Theta / \delta_0 = \delta_0; d, c) = c + (1 - c) \cdot [1 + \exp(-d(\Theta - \delta_0))]^{-1} \quad (2.13)$$

и основана на формуле полной вероятности двух событий. Одно из этих событий состоит в использовании возможности угадать правильный ответ на данное задание с вероятностью c даже в том случае, когда $\Theta \ll \delta_0$.

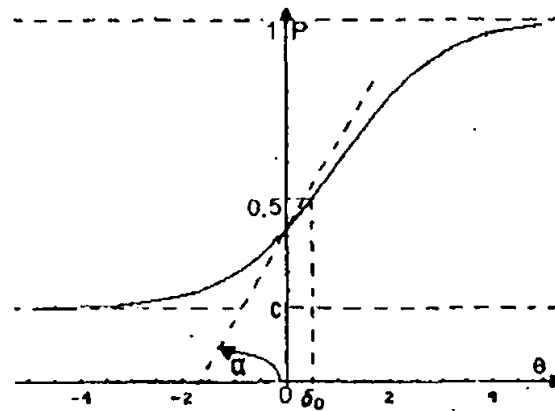


Рис. 2.3. Характеристическая кривая трудности δ_0 с дискриминационной возможностью $d=4 \operatorname{tg} \alpha$ и шансом угадывания c

Другое событие состоит в решении задания, и вероятность успеха при этом определяется формулой (2.10) в соответствии с уровнем подготовленности Θ участника тестирования. Очевидно, что предел функции (2.13) при $\Theta \rightarrow \infty$ равен c и потому ее график имеет горизонтальную асимптоту с уравнением $p=c$, см. рисунок 2.3. (на рис. 2.2. горизонтальная асимптота совпадает с c осью абсцисс $p=0$). Заметим, что функция (2.13) уже не является логистической. Можно ввести и большее количество параметров, более полно характеризующих тестовые задания и участников тестирования. Однако надежность их практического определения по результатам массового тестирования при этом заметно снижается.

2.3. Формирование матрицы ответов и первичных баллов

Количество участников тестирования различной подготовленности Θ_i , обозначим буквой n ; $i = 1, 2, \dots, n$. Каждому участнику предлагается один и тот же вариант теста, состоящий из k заданий различной трудности δ_j ; $j=1, 2, \dots, k$. Результат выполнения каждого задания оценивается по дихотомному принципу: ставится единица, если задание выполнено правильно, и ноль, если задание выполнено неверно. Множество всех таких единиц и нулей образует прямоугольную таблицу-матрицу размерностью n, k . Мы будем обозначать эту матрицу буквой $A=(a_{ij})$ и называть матрицей ответов. Она имеет n строк и k столбцов. Ее элемент a_{ij} , стоящий на пересечении i -ой строки и j -го столбца, выражает возможный результат выполнения i -м участником j -го задания. Элементы a_{ij} являются величинами случайными: они принимают значение 1 с вероятностью $p_{ij}=p(\Theta_i, \delta_j)$, и значение 0 с вероятностью $q_{ij}=1-p_{ij}$. При необходимости различать случайные величины от их реализации мы будем обозначать последние как $\bar{a}_{ij}$.

Матрица ответов представляет собой ту исходную информацию, по которой предполагается оценить латентные параметры тестирования - трудности заданий и уровни подготовленности испытуемых. Матрица имеет ярко выраженную вертикальную структуру - количество столбцов k (число зада-

ний в одном тесте) обычно не превышает 60-70, но количество строк n (число участников тестирования) имеет порядок, как правило, нескольких сотен или нескольких тысяч. Таким образом, матрица ответов содержит несколько десятков или даже сотен тысяч нулей и единиц. Однако такой громадный массив исходных данных часто удается существенно уменьшить (редуцировать) без потери информации. В основе такой редукции лежит понятие достаточной статистики. Вообще статистикой называется всякая функция результатов исходных наблюдений (в нашем случае элементов a_{ij} матрицы ответов A), проявляющая статистическую устойчивость в том смысле, что значение этой функции при повторных наблюдениях может быть предсказано с существенно лучшей точностью, чем результат отдельного наблюдения.

Удобными статистиками при обработке результатов массового тестирования являются частичные суммы элементов матрицы ответов A по каждой строке и по каждому столбцу, то есть числа

$$b_i \equiv a_i = \sum_{j=1}^k a_{ij}, \quad i = 1, 2, \dots, n \quad (2.14)$$

и

$$c_j \equiv a_j = \sum_{i=1}^n a_{ij}, \quad j = 1, 2, \dots, k \quad (2.15)$$

Число $b_i \equiv \bar{a}_i$ представляет собой количество верно выполненных заданий участником с номером i и представляет первичный балл i -го участника. Он отражает определенную и меру успеха i -го испытуемого при выполнении k заданий данного теста.

Первичные баллы, как и любые другие статистики, являются величинами случайными, но при повторных тестированиях они гораздо более устойчивы, чем случайные элементы матрицы ответов, т.е. результаты решения какого-нибудь одного задания каким-нибудь одним участником.

Случайные элементы a_{ij} матрицы ответов A являются индикаторами успешного решения i -м участником j -го задания, т.е.

$a_{ij} = \{1, \text{ если } i\text{-й участник верно решил } j\text{-ое задание,}$

$a_{ij} = \{0, \text{ если } i\text{-й участник неверно решил } j\text{-ое задание.}$

Поэтому математическим ожиданием a_{ij} является соответствующая вероятность p_{ij} , а дисперсия равна $p_{ij}q_{ij}$, т.е.

$$M\{a_{ij}\} = 1 \cdot p_{ij} + 0 \cdot q_{ij} = p_{ij} \quad (2.16)$$

$$D\{a_{ij}\} = M\{a_{ij}^2\} - M^2\{a_{ij}\} = 1^2 \cdot p_{ij} + 0^2 \cdot q_{ij} - p_{ij}^2 = p_{ij}q_{ij} \quad (2.17)$$

На этом основании можно найти математические ожидания и дисперсии первичных баллов. В самом деле,

$$M\{b_i\} = M\left\{\sum_{j=1}^k a_{ij}\right\} = \sum_{j=1}^k M\{a_{ij}\} = \sum_{j=1}^k p_{ij} = p_i \quad (2.18)$$

$$D\{b_i\} = D\left\{\sum_{j=1}^k a_{ij}\right\} = \sum_{j=1}^k D\{a_{ij}\} = \sum_{j=1}^k p_{ij} \cdot q_{ij} \quad (2.19)$$

Первичные баллы имеют обобщенное биномиальное распределение. Распределение соответствующих вероятностей $P_k(b_i)$ удобно описывать с помощью так называемой производящей функции вида [10]

$$\varphi_k(x) = (p_{i1}x + q_{i1})(p_{i2}x + q_{i2}) \dots (p_{ik}x + q_{ik}) \quad (2.20)$$

Вероятность $P_k(a_i)$ того, что участник с N_i^0 из k заданий верно решит b_i штук, $0 \leq b_i \leq k$, равно коэффициенту при x^{b_i} в разложении производящей функции (2.20) по степеням x . Здесь

$$p_{ij} = p(\Theta_i, \delta_i) \quad q_{ij} = 1 - p_{ij} \quad (2.21)$$

Например, если $k=2$, то

$$\varphi_{i2}(x) = (p_{i1}x + q_{i1})(p_{i2}x + q_{i2}) = p_{i1}p_{i2}x^2 + (p_{i1}q_{i2} + p_{i2}q_{i1})x + q_{i1}q_{i2}$$

Вероятность $P_{i2}(2)$ того, что i -ый участник решит оба задания верно, равна $p_{i1}p_{i2}$; коэффициент при x дает вероятность $P_{i2}(1)$ того, что верно решено лишь одно какое-нибудь задание; коэффициент при x^0 , т.е. свободный член $q_{i1}q_{i2}$ равен вероятности $P_{i2}(0)$ того, что оба задания решены ошибочно. Контролем служит то, что сумма коэффициентов при всевозможных степенях x равняется единице, то есть в нашем примере $P_{i2}(2) + P_{i2}(1) + P_{i2}(0) = 1$.

Вероятность того, что при наличии k заданий первичный балл участника с N^0 , окажется: а) менее m ; б) более m ; в) не менее m ; г) не более m , - можно найти по формулам, соответственно

$$\begin{aligned} &P_{ik}(0) + P_{ik}(1) + \dots + P_{ik}(m-1) \\ &P_{ik}(m+1) + P_{ik}(m+2) + \dots + P_{ik}(k) \\ &P_{ik}(m) + P_{ik}(m+1) + \dots + P_{ik}(k) \\ &P_{ik}(0) + P_{ik}(1) + \dots + P_{ik}(m) \end{aligned} \quad (2.22)$$

Если предположить, что все задания теста имеют одинаковый уровень трудности, т.е. $\delta_i = \delta$ для $\forall_i$ от 1 до k и, следовательно, $p_{ij} = p_i$, независимо от j , то формулы (2.18)-(2.20), конечно, упрощаются и принимают вид, соответственно [9]

$$M\{b_i\} = kp_i, \quad D\{b_i\} = kp_i q_i \quad (2.23)$$

$$\varphi_{ik}(x) = (p_i x + q_i)^k, \quad (2.24)$$

где

$$p_i = p(\Theta_i, \delta), \quad q_i = 1 - p_i$$

и коэффициентом при x^b , $0 \leq b \leq k$ в разложении бинома производящей функции (2.24) по степеням x служат биномиальные коэффициенты. Таким образом,

$$P_{ik}(b) = C_k^b \cdot p_i^b \cdot q_i^{k-b} \quad (2.25)$$

Здесь число сочетаний из k по b определяется формулой

$$C_k^b = \frac{k!}{b!(k-b)!}$$

Соотношение (2.25) называется формулой Бернулли, а соответствующее распределение вероятностей называется биномиальным. Тот факт, что к равнотрудным заданиям в тесте первичный балл b имеет биномиальное распределение будем записывать в виде

$$b \in B_i(k, p)$$

Таким образом, биномиальное распределение - это распределение числа "успехов" в k ; независимых испытаниях с двумя исходами ("успех" - "не-успех") и неизменной вероятностью "успеха" $p \in (0;1)$. В частности,

$$a_{ij} \in B_i(1, p_{ij}),$$

где a_{ij} - элемент матрицы ответов, $p_{ij} = p(\Theta_{ij}, \delta_j)$ - функция успеха.

Относительно формул (2.18), (2.19) и (2.23) заметим, что выражения для дисперсии справедливы только в предположении независимости случайных величин a_{ij} , а для справедливости формул для математического ожидания эта независимость не требуется.

На практике формулу (2.25) удобно заменить приближенным соотношением (замена тем точнее, чем больше k) вида (локальная теорема Лапласа) [9]

$$P_k(b) \approx \frac{1}{\sqrt{k p_i q_i}} \varphi(u) \quad (2.26)$$

где

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, \quad u = \frac{b - k p_i}{\sqrt{k p_i q_i}} \quad (2.27)$$

Функция $\varphi(u)$ представляет собой плотность распределения вероятностей центрированной (то есть имеющей нулевое математическое ожидание) и нормированной (то есть имеющей единичную дисперсию) случайной величины u , подчиняющейся нормальному закону, то есть $u \in N(0;1)$.

Вероятность попадания первичного балла b_i в заданный промежуток $[m_1; m_2]$ удобно прогнозировать не с помощью формул (2.22), а с помощью следующей интегральной теоремы Лапласа [9]:

$$P_k(m_1 \leq b_i \leq m_2) \approx \Phi(x_2) - \Phi(x_1) \quad (2.28)$$

Здесь

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du \quad (2.29)$$

- функция Лапласа.

$$x_1 = \frac{m_1 - kp_i}{\sqrt{kp_i q_i}}, \quad x_2 = \frac{m_2 - kp_i}{\sqrt{kp_i q_i}} \quad (2.30)$$

Функция $\varphi(x)$ четная, а функция $\Phi(x)$ нечетная.

Формулы (2.26), (2.28) удобно использовать для приближенного прогноза результатов тестирования.

Пример 1. Пусть тест содержит $k=20$ заданий различной трудности, а p обозначает усредненную вероятность правильного решения одного задания этого теста. Тогда участников тестирования с различным уровнем подготовленности Θ_i можно классифицировать по величине соответствующей вероятности $p_i = p(\Theta_i, \delta)$. Будем условно называть уровень подготовленности очень слабым, если $p_i < 0,3$; слабым, если $0,3 \leq p_i \leq 0,55$; средним, если $0,55 < p_i \leq 0,75$; хорошим, если $p_i > 0,75$.

Тест имеет закрытую форму, если его задания сопровождаются несколькими (обычно пятью) вариантами ответа, из которых только один является верным, а остальные лишь правдоподобны. В таком случае вероятность $p=0,2$ соответствует бездумному угадыванию ответа. Получить при этом первичный балл выше 7-8 практически не представляется возможным.

Первичные баллы в результате тестирования являются, как уже отмечалось выше, величинами случайными. При реальном количестве числа k заданий в тесте их распределение приближенно нормально. Поэтому доверительный интервал для них можно построить с помощью следующего соотношения:

$$P_y \{ |b - M(b)| \leq \varepsilon \} = 2\Phi\left(\frac{\varepsilon}{\sqrt{D(b)}}\right) \quad (2.31)$$

Здесь математическое ожидание $M(b)$ и дисперсия $D(b)$ определяются формулами (2.18), (2.19) или (2.23) ε - любая положительная константа, определяющая длину доверительного интервала 2ε ; Φ - функция Лапласа (2.29).

Более точный результат можно получить с помощью следующего факта:

$$P_{ik} \{ m_1 < M(b_i) < m_2 \} = d_i \quad (2.32)$$

Пример 2. Некоторый участник тестирования с N_i^0 при выполнении теста, состоящего из $k=20$ заданий, набрал $\bar{b}_i=16$ баллов. Какова вероятность d того, что реальное значение первичного балла находится на отрезке $[14;18]$?

Для ответа на этот вопрос воспользуемся сначала простейшей схемой (2.14),(2.31), приближенно заменив нужную для этого вероятность p_b - ее оценкой $\bar{p}_i = \frac{16}{20} = 0,8$ и взяв $\varepsilon=2$. Имеем

$$2\Phi\left(\frac{2}{\sqrt{20 \cdot 0,8 \cdot 0,2}}\right) = 2\Phi(1,12) = 0,74 = d$$

Теперь поставим вопрос иначе: каков доверительный интервал для $\bar{b}_i = 16$ с уровнем значимости $d=0,74$? Пользуясь формулами (2.33), имеем

$$m_1 = \frac{20}{20 + 1,12^2} \left(16 + 0,63 - 1,12 \sqrt{\frac{16 \cdot 4}{20} + 0,63} \right) = 0,94(16,63 - 2,19) = 13,6;$$

$$m_2 = 0,94(16,63 + 2,19) = 17,7$$

Округляя m_1 и m_2 до целых значений, получаем доверительный промежуток $[14;18]$.

Все формулы данного раздела, начиная с (2.18), и оба примера составлены применительно к первичным баллам участников (2.14). Однако все остается верным и применительно к первичным баллам заданий (2.15). Достаточно лишь в формулах (2.18)-(2.33) заменить b_i на c_j , k на n и p^i на p_j .

Пример3. В тестировании, в котором приняли участие $n=100$ человек, первичный балл задания с N_j^0 оказался равным $\bar{c}_j = 80$. Каков доверительный интервал для этого первичного балла с уровнем значимости $d=0,95$?

Концы искомого интервала получим по формулам (2.33), которые надо понимать следующим образом:

$$\begin{aligned}
 m_1 &= \frac{n}{n+l^2} \left(c_i + \frac{l^2}{2} - l \sqrt{\frac{\bar{c}_j (n - \bar{c}_j)}{n} + \frac{l^2}{2}} \right); \\
 m_2 &= \frac{n}{n+l^2} \left(\bar{c}_j + \frac{l^2}{2} + l \sqrt{\frac{\bar{c}_j (n - \bar{c}_j)}{n} + \frac{l^2}{2}} \right);
 \end{aligned}
 \tag{2.34}$$

Имеем $l = 1.96 \Rightarrow m_1 \approx 71, \quad m_2 \approx 87$.

В заключение раздела сделаем два замечания [24].

Замечание 1. Формулы (2.19) и (2.23) для дисперсии первичных баллов справедливы лишь в том случае, когда случайные величины a_{ij} независимы, т.е. вероятности решения, например, i -м участником различных заданий не зависят друг от друга. В общем случае дисперсия определяется формулой

$$D\{b_i\} = \sum_{j=1}^k D\{a_{ij}\} + 2 \sum_{j < l} k_{i,jl} \tag{2.35}$$

где $k_{i,jl}$ — ковариационный момент случайных величин a_{ij} и a_{il}

$$k_{i,jl} = M\{a_{ij} \cdot a_{il}\} - M\{a_{ij}\} \cdot M\{a_{il}\} \tag{2.36}$$

Величина $a_{ij} \cdot a_{il}$ равна единице только в том случае, когда i -ый участник решил j -ое, и l -ое задания. Обозначим вероятность этого события $p_{i,jl}$. Тогда [25]

$$M\{a_{ij} \cdot a_{il}\} = p_{i,jl}, \quad k_{i,jl} = p_{i,jl} - p_{ij} \cdot p_{il} \tag{2.37}$$

Поэтому

$$D\{b_i\} = \sum_{j=1}^k p_{ij} \cdot q_{ij} + 2 \sum_{j < l} (p_{i,jl} - p_{ij} \cdot p_{il}) \tag{2.38}$$

и, таким образом, для вычисления дисперсии первичного балла недостаточно знать только вероятности p_{ij} ; правильного решения i -м участником j -го задания: надо еще звать вероятности $p_{i,jl}$ совместного решения каждой пары j, l заданий.

В частности, если $\forall_i p_{i,jl}$ не зависит от j и l , то формула (2.38) упрощается и принимает вид

$$D\{b_i\} = kp_i q_i + k(k-1)(p_i^0 - p_i^2) \tag{2.39}$$

где p — вероятность правильного решения i -м участником любой пары заданий.

Такая замена не только практически удобна, но и в определенном смысле и теоретически оправдана [15]. В самом деле, математические ожидания первичного балла, вычисленные по формулам (2.18) совпадают, поскольку

$$\sum_{j=1}^k p_{ij} = p_i k \quad (2.40)$$

Важным свойством первичных баллов является то, что в рамках основной модели Раша они являются статистиками достаточными. Прежде, чем обосновывать и интерпретировать этот факт, приведем соответствующие короткие справочные сведения из математической статистики.

2.4. Формирование достаточных статистик оценки уровня знаний

Достаточной статистикой для семейства распределений вероятностей $\{P_\lambda; \lambda \in \Lambda\}$ или для параметра (в общем случае, векторного) $\lambda \in \Lambda$ называется такая статистика X (векторная случайная величина), что для любого события S существует вариант условной вероятности $P_\lambda(X=x)$, не зависящий от λ . Это эквивалентно требованию, что условное распределение любой другой статистики Y при условии $X=x$ не зависит от λ [25].

Статистика X достаточна для семейства $\{P_\lambda\}$ в том и только в том случае, когда соответствующая плотность распределения p_λ может быть факторизована (разложена на множители) в виде :

$$p_\lambda(\omega) = g_\lambda(X(\omega)) \cdot h(\omega), \quad (2.41)$$

где g_λ и h - неотрицательные функции события ω , причем g_λ зависит от λ , но от ω зависит только через $X(\omega)$, тогда как h не зависит от λ .

Пример 1. Участникам тестирования предложено k заданий одинаковой трудности. Уровень подготовленности участника предполагается оценивать вероятностью p правильного решения каждого задания. Каждому участнику в матрице ответов соответствует одна строка, представляющая собой последовательность независимых случайных величин, принимающих значение 1 с неизвестной вероятностью p и значение 0 с вероятностью $1-p$ (схема Бернул-

ли). Вероятность конкретной реализации $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$ указанной последовательности определяется теоремой умножения вероятностей и имеет вид

$$p_p(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k) = \prod_{j=1}^k p^{\bar{a}_j} (1-p)^{1-\bar{a}_j} = p^{\sum_{j=1}^k \bar{a}_j} \cdot (1-p)^{k - \sum_{j=1}^k \bar{a}_j} \quad (2.42)$$

Равенство (2.41) выполняется, если положить

$$X = \sum_{j=1}^k a_j, \lambda = p, \quad g_k = p_p, \quad h=1$$

и потому первичные баллы ; участников являются достаточными статистиками - вероятность (2.42) зависит от результатов тестирования $\bar{a}_1, \dots, \bar{a}_k$ только через маргинальную сумму $X = \sum_{j=1}^k a_j$. В четности, эмпирическая частота

$\bar{p} = \frac{1}{k} \sum_{j=1}^k a_j$ является достаточной оценкой для искомой вероятности p .

Пример 2. Пусть уровни подготовленности n участников тестирования составляют последовательность независимых нормально распределенных величин $\Theta_1, \Theta_2, \dots, \Theta_n$ с неизвестными средними значениями μ и дисперсией σ^2 . Совместная плотность распределения имеет вид

$$\begin{aligned} p_{p, \sigma^2}(\tilde{\Theta}_1, \tilde{\Theta}_2, \dots, \tilde{\Theta}_n) &= (2\pi\sigma^2)^{-n/2} \cdot \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (\tilde{\Theta}_i - \mu)^2\right] = \\ &= (2\pi\sigma^2)^{-n/2} \cdot \exp\left[-\frac{n\mu^2}{2\sigma^2} - \frac{1}{2\sigma^2} \sum_{i=1}^n \tilde{\Theta}_i^2 + \frac{\mu}{\sigma^2} \sum_{i=1}^n \tilde{\Theta}_i\right] \end{aligned} \quad (2.43)$$

который зависит от $\Theta_1, \dots, \Theta_n$ только через величины $\sum \tilde{\Theta}_i$ и $\sum \tilde{\Theta}_i^2$. Поэтому векторная статистика

$$X = \left(\sum_{i=1}^n \tilde{\Theta}_i, \sum_{i=1}^n \tilde{\Theta}_i^2 \right)$$

является достаточной статистикой для двумерного параметра $\lambda = (\mu, \sigma^2)$. Достаточными оценками среднего μ и дисперсии σ^2 служат функции

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \tilde{\Theta}_i, \quad \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (\tilde{\Theta}_i - \hat{\mu})^2 \quad (2.44)$$

которые могут быть выражены через элементы вектора X .

Знание достаточной статистики X даст исчерпывающий материал относительно параметра λ , поскольку любые дополнительные статистические данные ничего не добавляют к той информации о параметре, которая содержится в распределении X . Переход от исходного семейства распределений к семейству распределений достаточной статистики называется редукцией статистической задачи. Смысл редукции состоит в уменьшении (часто весьма значительном) количества исходных данных без потери содержащейся в этих данных информации.

2.5. Редукция матрицы ответов

В рамках основной модели Раша первичные баллы являются достаточными статистиками [39]. Вероятность того, что участник тестирования с N_i^0 правильно решит задание с N_j^0 , следовательно, (i, j) —позицию в матрице ответов A займет 1, определяется формулой

$$p_{ij} = \frac{s_i}{s_i + t_j} = \frac{s_i / t_j}{1 + s_i / t_j} = \frac{\xi_{ij}}{1 + \xi_{ij}} \quad (2.45)$$

где s_i — уровень подготовленности i -го участника, t_j - уровень трудности j -го задания, а $\xi_{ij} = s_i / t_j$

Вероятность неверного решения и, следовательно, равенства a_{ij} нулю равна:

$$q_{ij} = 1 - p_{ij} = 1 - \frac{s_i / t_j}{1 + s_i / t_j} = \frac{1}{1 + s_i / t_j} = \frac{1}{1 + \xi_{ij}}$$

Выражения для p_{ij} и q_{ij} можно записать единой формулой

$$\frac{(\xi_{ij})^{a_{ij}}}{1 + \xi_{ij}} = \frac{(s_i / t_j)^{a_{ij}}}{1 + s_i / t_j} = \begin{cases} p_{ij}, & \text{если } a_{ij} = 1 \\ q_{ij}, & \text{если } a_{ij} = 0 \end{cases} \quad (2.46)$$

Поэтому вероятность конкретной реализации $\tilde{A} = (a_{ij})$ матрицы ответов, определяемая теоремой умножения вероятностей, имеет вид

$$\begin{aligned}
 p_{s,t}(\tilde{A}_{n,k}) &= \prod_{i=1}^n \prod_{j=1}^k \frac{(s_i / t_j)^{\tilde{a}_{ij}}}{1 + s_i / t_j} = \frac{\prod_{i=1}^n \prod_{j=1}^k s_i^{\tilde{a}_{ij}} \cdot \prod_{i=1}^n \prod_{j=1}^k t_j^{-\tilde{a}_{ij}}}{\prod_{i=1}^n \prod_{j=1}^k (1 + s_i / t_j)} = \\
 &= \frac{\prod_{i=1}^n s_i^{\bar{b}_i} \prod_{j=1}^k \prod_{i=1}^n t_j^{-\tilde{a}_{ij}}}{\prod_{i=1}^n \prod_{j=1}^k (1 + s_i / t_j)} = \frac{\prod_{i=1}^n s_i^{\bar{b}_i} \prod_{j=1}^k t_j^{-\bar{c}_j}}{\prod_{i=1}^n \prod_{j=1}^k (1 + s_i / t_j)}
 \end{aligned} \tag{2.47}$$

Последнее выражение показывает, что рассматриваемая вероятность зависит от элементов a_{ij} матрицы ответов A только через ее маргинальные суммы $\bar{b}_i = \bar{a}_i, \bar{c}_j = \bar{a}_j$ по строкам и по столбцам. Поэтому первичные баллы участников тестирования и заданий теста являются достаточными статистиками для искомых параметров s_i и t_j соответственно ($i = 1, 2, \dots, n; j = 1, 2, \dots, k$). Следовательно, при оценивании параметров можно вместо $n \cdot k$ элементов матрицы ответов оперировать лишь с первичными баллами, общее количество которых равно $n + k$. При этом количество исходных чисел уменьшается в десятки раз без какой-либо потери информации.

Другой вывод, который вытекает из достаточности первичных баллов, состоит в следующем. Статистические оценки уровней подготовленности всех участников тестирования, набравших одинаковое количество первичных баллов, совпадают, поскольку являются функциями равных достаточных статистик. То же можно сказать и относительно оценок уровней трудности заданий - они совпадают для всех заданий, характеризующихся одинаковыми первичными баллами. Таким образом, не зная реальных оценок параметров s и t и даже не обсуждая методы их вычисления, мы можем упорядочить всех участников тестирования в порядке возрастания оценок их уровней подготовленности только на основании первичных баллов участников. Аналогично, можно упорядочить все задания теста в порядке возрастания оценок их трудностей только на основании первичных баллов заданий. Необходимо только иметь в виду, что речь идет об упорядочивании именно статистических оценок латентных параметров, а не их истинных значений, которые

остаются недоступными и располагаются, возможно, в ином порядке, чем их оценки. В большей степени это относится к характеристикам участников тестирования, чем к характеристикам заданий, поскольку точность соответствия оценок параметров (величин случайных) их истинным значениям (константам) зависит от объема исходных данных, доступных для вычисления оценок. Поэтому проявляется тот факт, что первичные баллы участников выводятся на основании k исходных элементов матрицы ответов (маргинальные суммы по строкам), а при вычислении первичных баллов заданий используется n таких элементов (маргинальные суммы по столбцам), причем, как уже отмечалось выше, n , как правило, значительно превосходит k .

Предположим, например, что тест, состоящий из 20 заданий, выполняют 1000 участников тестирования. Тогда $n=1000$, $k=20$. Поэтому первичные баллы каждого из 1000 участников могут принимать только ограниченное число (21) дискретных значений $0, 1, 2, \dots, 19, 20$. В то же время, количество возможных значений параметра z бесконечно, и они непрерывно заполняют, вообще говоря, всю положительную полуось. Из этих простых рассуждений ясно, что участники тестирования, набравшие одинаковое количество первичных баллов, почти наверняка имеют различные истинные уровни подготовленности.

2.6. Разработка моделей преобразования балльных шкал

С математической точки зрения, процесс измерения есть отображение состояния измеряемого объекта на некоторое множество действительных чисел (или на некоторое множество точек числовой оси), называемое шкалой. Однако шкала - это не просто определенное множество - дискретное или сплошь заполняющее некоторый промежуток. Важнейшей отличительной чертой шкалы является набор тех соотношений между ее элементами (отсчетами), которые имеют содержательный смысл и разумное толкование в рамках этой шкалы [26,29].

Существует много разлйных шкал, в том числе, в педагогике:

1 Порядковые шкалы, где результаты измерений осмысленно можно только сравнивать между собой. Примером может служить принятая в школе система оценок, выставляемых ученикам в зависимости от их успехов в учебе. Из соотношения отметок $b_1 < b_2$ для учеников A_1 и A_2 можно лишь заключить, что A_1 учится хуже A_2 . Если же, например, $b_2 - b_1 = 1$, то утверждение "успехи A_2 на 1 выше, чем успехи A_1 " не объясняет, каково различие между учениками и, по существу, лишено смысла. То же можно сказать и относительно шкалы первичных баллов (в абсолютном или относительном выражении) как для участников тестирования, так и для тестовых заданий.

2. Шкала более высокого уровня называется интервальной, или метрической.

Ее отличительной чертой является наличие метрики. Это означает, что для любых отсчетов b_1 и b_2 содержательный смысл имеют не только соотношения типа $b_1 < b_2$ или $b_1 = b_2$, но и разность $b_2 - b_1$. При этом $|b_2 - b_1|$ трактуется как расстояние (между двумя элементами метрического множества), выраженное в определенных единицах и, главное, имеющее осмысленное толкование. Специфика шкалы состоит в отсутствии нулевого штриха, то есть в отсутствии начала отсчета. Поэтому метрическая шкала прекрасно подходит для фиксации взаимного положения измеряемых объектов (относительно друг друга), но она не в состоянии информировать о местоположении объекта в некоторой единой системе координат (удалении от начала отсчета). С математической точки зрения указанная ситуация означает, что на множестве определена метрика, единица измерения расстояния, но нет понятия нормы (определено понятие "расстояние", но нет понятия "длина"). Например, при строительстве гидросооружений важно измерять превышения (разности высот) между определенными точками (взаимное положение по высоте, имеющее конкретную гидродинамическую трактовку), но не сами высоты. Превышение между двумя точками, имеющие высоты, например 48 м и 45 м, имеет то же смысл, что и превышение между точками с высотами 5 м и 2 м. В противоположность этому разности между первичными баллами 48-45 и 5-2

невозможно сравнивать осмысленно.

По такой же, по существу, шкале, по которой измеряются превышения, измеряются и латентные параметры Θ и δ , но только единицей измерения расстояний служат не метры, а логиты.

Таблица 2.2.

Разность $\Theta-\delta$	Вероятность P верного решения	Информация в ответе $=pq$	Относительная эффективность в процентах
5	0,99	0,01	4
4	0,98	0,02	8
3	0,95	0,05	20
2	0,88	0,11	44
1	0,73	0,20	80
0	0,50	0,25	100
-1	0,27	0,20	80
-2	0,12	0,11	44
-3	0,05	0,05	20
-4	0,02	0,02	8
-5	0,01	0,01	4

Таблица 2.2. указывает соотношение между разностями $\Theta-\delta$ в логитах и их трактовкой в виде вероятности (2.6) того, что задание трудности δ будет верно выполнено участником с уровнем подготовленности Θ . Данные этой таблицы никак не изменятся, если к величинам Θ и δ прибавить любую константу. Последний столбец таблицы 2.2. содержит произведения $p(1-p)$, которые можно трактовать как количество информации о разности $\Theta-\delta$, которое содержится в соответствующем элементе матрицы ответов. Содержание этого столбца мы используем позже, но уже сейчас полезно отметить, что информативность ответов зависит только от расстояния $|\Theta-\delta|$ между Θ и δ и заметно падает с увеличением этого расстояния. Так, одно задание максимальной эффективности равносильно (с точки зрения поддержания одной и той же точности измерения) около 25 заданий минимальной эффективности.

Если признаков, различающих объекты или события, только два, то номинальная шкала называется дихотомной. Примером могут служить элементы матрицы ответов участников тестирования на задания теста: правильное выполнение задания ("да") обозначается единицей, ошибочное ("нет") - нулем. При этом разность 1-0 не имеет никакого смысла, и сами цифры 1 и 0 можно заменить любыми другими, например, цифрами 9 и 5, символами "+", "-" словами "да", "нет", "зачет", "незачет" и т.п.

Соответствующие номинальным шкалам данные состоят из наблюдаемых значений частот или табличных сведений о числе появлений каждой из разновидностей изучаемой переменной. Для характеристики номинальных данных часто используются такие (дескриптивные) статистики, как пропорция и процентное отношение.

Использование той или иной шкалы из перечисленных четырех накладывает отпечаток и на применимость тех или иных методов математической обработки, которой обычно подвергаются исходные данные. Например, регрессионный анализ применим только по отношению к количественно выраженным переменным, измеряемым, по крайней мере, в метрической шкале. Примерно то же самое можно сказать и относительно наиболее известных методов корреляционного анализа. Сказанное не означает, что результаты тестирования, отнесенные к порядковым или даже номинальным шкалам, нельзя анализировать количественно. Однако методы такого анализа должны быть, в общем случае, специальными и отличными от тех, которые используются для переменных в шкалах метрических нормированных. Например, даже такую общепринятую оценку центра рассеяния переменной как арифметическая средняя часто бывает более обоснованным заменить медианой вариационного ряда, если переменные отнесены к шкале порядковой, а не метрической.

Выводы по главе 2

1. Показано, что в метрической шкале определено не только понятие метрики, но и понятие нормы, позволяющее измерять "длины" (то есть определенное местоположение относительно нуля, начала отсчета). Поэтому в такой шкале имеет смысл говорить не только о разностях типа $\Theta - \delta$, но и о каждой величине Θ или δ в отдельности.

2. Показано, что содержательная интерпретация результатов математической обработки данных тестирования может быть дана лишь в том случае, если методы этой обработки адекватны тем шкалам, к которым отнесена исходная информация.

3. Предложен подход к оценке результатов ответа на УТЗ, в котором предлагается использовать расширение модели адаптивного тестового контроля, где сложность следующего задания выбирается на основании решения предыдущего. При этом неточный ответ на УТЗ определенной сложности можно рассматривать как решение УТЗ пониженной сложности.

3. РАЗРАБОТКА ПРИНЦИПОВ КОНСТРУИРОВАНИЯ ТЕСТОВЫХ ЗАДАНИЙ И ФОРМАЛИЗАЦИЯ КОМПОНЕТОВ СИСТЕМЫ АДАПТИВНОГО ТЕСТОВОГО КОНТРОЛЯ

В третьей главе ставятся и решаются задачи формализация компонентов системы тестового контроля. Разработаны принципы формирования тестовых заданий и сформулированы требования к программной среде конструирования тестовых заданий.

3.1. Построение формализованной модели тестового контроля

Система тестового контроля является неотъемлемой частью системы переподготовки. Она должна иметь развитые средства администрирования, конструирования тестовых заданий, формирования алгоритмических структур тестов и их предъявления. Кроме того, объединение форм конструирования тестовых заданий с адаптивными механизмами их предъявления даст неоспоримое преимущество системе тестового контроля.

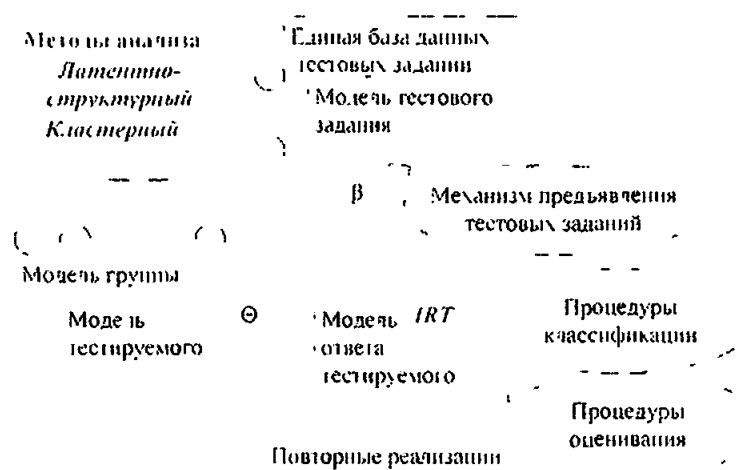


Рис. 3.1. Система моделирования тестирования

С точки зрения конструирования тестового задания предлагается стандартизация формы представления в графическом формате, что позволит

использовать различные общесистемные платформы. В результате проведенных исследований показано, что все наиболее распространенные формы, а именно, «закрытая», «на соответствие», «на упорядочение» сводятся к одной графической форме «на перетаскивание». На основе созданного описания тестового задания конструктор тестовых заданий создает «Flash» - файл, который можно включить в HTML и использовать для режимов тренажера и самоконтроля. Наиболее ценной функциональной возможностью конструктора является динамическая генерация «Flash» - файлов тестовых заданий с регистрацией ответов. В силу того, что основной задачей «Flash» - технологий является их эффективность в Интернет, то подсистема предъявления тестовых заданий и динамического конструирования адаптивных тестов обладает той же эффективностью.

Конструктор должен иметь качественные формы представления и удобный интерфейс формирования тестовых заданий, что существенно влияет на время подготовки тестов. Критерий времени важен, так как количество заданий должно быть велико, а содержательная часть заданий по многим предметным областям очень быстро устаревает. Проигрыватель тестовых заданий должен работать как в локальном, в сетевом варианте, так и Интернет вариантах.

Наиболее эффективными являются адаптивные алгоритмы предъявления заданий с учетом их сложности, поэтому часть диссертационных исследований направлена на разработку методов экспертного оценивания уровня сложности заданий с учетом их последующей корректировки по результатам статистической обработки ответов тестируемых. Включение математических моделей тестируемого и эффективности процедур компьютерного тестового контроля в системы переподготовки позволит моделировать образовательную траекторию с точки зрения эффективности накопления информации, предъявляемой последовательности модулей.

В общем случае следует разделять механизм предъявления тестовых заданий и процедуры оценивания или классификации уровня знаний.

Механизм предъявления может быть построен различными способами

Статический когда количество заданий различной степени сложности, включенных в тест определено заранее. Вариации общим количеством заданий и их распределением по сложности дают различные результаты по точности оценивания уровня знаний, поэтому необходимо решение задач распределения в зависимости от априорной информации об уровне знаний

Адаптивный механизм предполагает выбор каждого последующего задания из базы тестовых заданий в зависимости от ответов на все предыдущие. В данном случае формирование алгоритма для механизма предъявления дает еще большую свободу в творчестве преподавателя

Механизм предъявления дает лишь последовательность заданий различной сложности. Оценивание уровня знаний тестируемого является следующей задачей. В данном случае имеет место два подхода классификация, когда количество значений оценок тестируемого определено заранее, и оценивание когда численное значение оценки может быть произвольным

Процедуры классификации могут базироваться на использовании теоретических положений *дискриминантного анализа*. Модель дискриминантного анализа при использовании в процедуре тестирования концептуально подобна использованию бальной шкалы, поэтому она позволяет провести анализ оценивания с заданием весов тестовых заданий различной степени сложности. Теоретические положения дискриминантного анализа позволяют оценить влияние коррелированности ответов, что не позволяет классическая теория

При моделировании эффективности процедуры в данном случае наиболее информативной характеристикой является матрица ошибочной классификации. Получение численных оценок значений матрицы в некоторых случаях возможно аналитическим способом, однако, в общем случае, необходимо использование *имитационного подхода*.

Разработана *имитационная модель* позволяющая вычислять значение ошибочной классификации для любого механизма предъявления. Особенностью модели является возможность сравнительного анализа рассматриваемых механизмов и процедур с любой степенью статистической значимости.

На рис.3.2. приведена формализованная схема этапов и компонент тестирования, где:

X - состояние среды, влияющей на результат тестирования;

X' - информация о среде X , получаемая системой тестирования;

Y - состояние тестируемого, определяемое датчиком D_Y , на выходе которого имеется Y' информация правильности ответов (об этом состоянии информация поступает в систему в виде ответов на вопросы Y).

Z^* - цели тестирования;

R - ресурсы, которыми система располагает для тестирования.

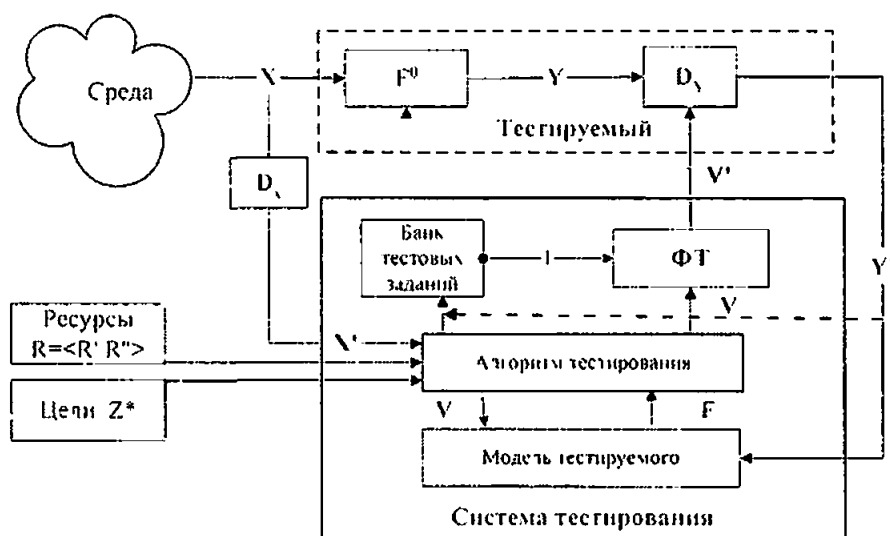


Рис. 3.2. Этапы и компоненты тестирования

Основная задача состоит в следующем: организовать процесс тестирования V , оценивающий способности Y субъекта таким образом, чтобы выполнялись поставленные цели Z^* .

$$V = A_1(X, Y, Z^*, R)$$

где A_T - алгоритм тестирования;

$R=\langle R', R'' \rangle$ - ресурс тестирования;

R' , - внешний ресурс, определяемый возможностями системы тестирования;

R'' - внутренний ресурс, выделяемый тестируемым;

F^0 - оператор, конкретного тестируемого.

Формулировка целей тестирования связана с определением критериев и требований к ним, выполнение которых решает задачу идентификации качества обучения:

$$Z^* : \{ \eta_i^{\geq}(s) \geq a_i \mid i=1, \dots, k_1 \mid \eta_j^{\equiv}(s) = b_j \mid j=1, \dots, k_2 \mid \eta_l^{\rightarrow}(s) \rightarrow \text{extr} \mid l=1, \dots, k_3 \}, \quad (3.1)$$

$\eta_i^{\geq}$, $\eta_j^{\equiv}$ и $\eta_l^{\rightarrow}$ - критерии (функционалы), определяемые на S - состояние объекта обучения и его среды: $S=\langle X, Y \rangle$

Цели-неравенства ($\eta_i^{\geq}$) определяют *min* знаний, умений и навыков обучаемого, нарушение которых недопустимо. Функционал φ_i может выражать в виде бальной оценки уровень знаний по i -ой учебной дисциплине. Тогда $a_i=3$.

Цели-равенства ($\eta_j^{\equiv}$) связаны с теми же знаниями, умениями и навыками обучаемого, отсутствие которых недопустимо (ядро знаний по учебной дисциплине). Это основные законы предметной области. Тогда $b_j=1$ определяет знание, а $b_j=0$ - незнание этого закона, что недопустимо, т.к. $b_j=1$.

Экстремальные цели ($\eta_l^{\rightarrow}$) связаны с теми же качествами обучаемого, которые целесообразно экстремизировать при обязательном выполнении предыдущих целей. Они определяют качество тестирования. Если $\eta_l^{\rightarrow}$ - средний балл по i -му предмету, то естественно требовать, чтобы $\eta_l^{\rightarrow} \rightarrow \text{max}$. Если $\eta_l^{\rightarrow}$ - время тестирования по i -му предмету, то $\eta_l^{\rightarrow} \rightarrow \text{min}$. Индивидуализация процесса обучения ориентирована именно на достижение экстремальных целей.

Цели Z^* обучения и ресурс, выделенный для достижения этих целей, определяют объект обучения, включающий вместе с учеником и его среду.

Субъектом в системе обучения является заказчик, которому предстоит использовать выпускников. Цели обучения представляют собой требования, предъявляемые квалификационной характеристикой к выпускнику.

Структурный синтез модели обучаемого связан с определением структуры модели F , связывающей состояние среды и тестирующего воздействия V , $Y=F(X, V)$. Под общей структурой будем понимать:

структуру $X=(x_1, \dots, x_n)$,

структуру $V=(v_1, \dots, v_q)$,

структуру $Y=(y_1, \dots, y_q)$,

структуру оператора $F=<W, C>$,

где W - структура и $C=(c_1, \dots, c_2)$ - параметры модели.

Описание модели обучаемого имеет вид:

$Y=f_1(X, V, C)$ где f_1 - выбранный оператор, параметры которого C предстоит определить далее. Надо помнить, что Y - характеристики состояния обучаемого, а Y' - ответы обучаемого на тестовые задания.

$Y'=f_2(X', Y, C)$ где f_2 - ответы обучаемого в измененной ситуации X' , которая определяется X . Влияние среды учитывается вопросами, задаваемому обучаемому по поводу его среды (прожиточный $\min$ родителей, условия его жизни и т.д.).

$X'=f_3(X, V')$, где f_3 - оператор, определяющий ответ обучаемого на вопрос V' по поводу источника среды X .

Полагаем, что результат тестирования используется для коррекции знаний обучаемого, т.е. что V' изменяет при необходимости состояние обучаемого. Кроме того, считаем, что алгоритм синтеза тестов и алгоритм обучения равны, т.е. $A_V = A_L$, где A_L - алгоритм обучения.

Синтез обучения заключается в определении порции информации U^* , обучение которой дает возможность добиться цели обучения.

$$Z^* : \{ \eta^2_i(x, y) \geq a_i \mid i=1, \dots, k_1 \mid \eta^2_j(x, y) = b_j \mid j=1, \dots, k_2 \mid \eta^2_l(x, y) \rightarrow extr \mid l=1, \dots, k_3 \}. \quad (3.2)$$

Подставляя сюда $Y=F(X, V)$, приходим к оптимизационной задаче

$$\eta^{\rightarrow}_l(X', F(X', V)) \rightarrow \text{extr} \quad l=1, \dots, k_3 \quad V \in \Omega, \quad (3.3)$$

где множество допустимого обучения представляет:

$$\Omega: \{ \eta^{\rightarrow}_l(X', F(X', V)) \geq a_l \mid \eta^{\rightarrow}_j(X', F(X', V)) = b_j \quad V \in R \}. \quad (3.4)$$

Для решения этой задачи следует осуществить свертку целей (3.3) в виде:

$$Q(X', V) = \sum_{l=1}^{k_3} \beta_l \eta_l(X', F(X', V)), \quad (3.5)$$

где $\beta_l > 0$ - вес цели.

В результате получим стандартную вариационную задачу

$$Q(X', V) \rightarrow \min \Rightarrow U^*, \quad (3.6)$$

где U^* - оптимальная порция обучающей информации для конкретного обучаемого.

Если целевые критерии являются функциями, то задача является задачей математического программирования.

Этап адаптации призван откорректировать систему тестирования в связи с ответами обучаемого (валидность, надежность, характеристикой теста). Модель обучаемого описывает оценку $\hat{Y}$ состояния Y , причем $\hat{Y} = F(X', V)$.

Алгоритм обучения определяет, чему следует учить, т.е. обучающее воздействие $U = A_V(X', \hat{Y}, Z^*, R)$. В простейшем случае $V=U$. Кроме того, A_V определяет тесты, ответы на которые несут информацию о модели F $V = \psi(X', \hat{Y})$, где A_V - алгоритм синтеза теста V .

Формирователь тестов (ФТ) работает аналогично: $V' = \psi(V, I)$, где I - набор сведений, необходимых для усвоения. Информация о состоянии обучаемого $Y' = D_V(Y, V')$. Здесь V - адрес теста; V' - содержание теста.

Важно отметить, что алгоритм обучения определяет порядок генерации и уровень сложности тестовых заданий.

уровня знаний в классе нормального распределения.

3.2. Принципы разработки тестовых заданий и формирование требований к инструментальной среде их конструирования

Следующая задача, которая решена в диссертации – это формализованное представление механизма конструирования тестовых заданий. Разработаны принципы создания инструментальной программной среды формирования тестовых заданий.

3.2.1. Разработка методических правил проектирования тестовых заданий

Применение методологических правил для проектирования тестовых ситуаций в области точных наук позволяет избежать конструирования некорректных тестовых заданий и указывает на возможность использования многообразия форм тестовых заданий при одном и то же содержании. Методологические правила конструирования теста в единстве с направленностью специальных знаний находят выражение в философии компьютерного адаптивного тестирования [33, 47].

П1: результаты интеллектуальной деятельности обучаемых, рассматриваемые как понимание, выражаются языковыми структурами, которыми они одинаково пользуются.

Содержание тестового задания, которое "видят" за словами на экране, должно восприниматься ими одинаково. Под объектом компьютерного адаптивного тестирования подразумеваются все сотрудники и рабочие предприятий, участвующие в аттестации, и уровень обученности которых априорно не установлен. Образ составляют тестируемые только одного уровня достижений (например, только "отличники"). Каждый конкретный тестируемый из заданного образа называется испытуемым.

Следующее правило П2 связано с необходимостью представления тестового задания в виде утверждения или повелительного предложения, из

которых, в зависимости от качества ответов, получается истинное или ложное высказывание.

П2: тестовое задание формулируется так, что в его структуре заложена возможность однозначного ответа, преобразующего это утверждение в истинное высказывание.

Тестовые задания не только строятся на основе и с помощью синтаксических, семантических и прагматических правил, но сами *содержат в себе задание и ответ*. С этой точки зрения тестового задания представляют собой когнитивные структуры, предназначенные для получения истинных утверждений в случае формулирования испытуемыми прагматически корректных ответов.

П3: тестовое задание конструируется с учетом момента времени накопленных обучаемым знаний в определенной предметной области.

Ставить вопрос в один ряд с тестовым заданием можно только в том случае, когда вопрос адресативен, причем отнесенность вопроса к самому себе - аномалия. Если считать требование ответа для каждого вопроса обязательным, необходимо иметь в виду, что при известных обстоятельствах отсутствие ответа - это тоже ответ. Отсюда следует правило П4:

П4: вопрос, который может предполагать отсутствие ответа, не может использоваться при конструировании тестового утверждения, так как по своим прагматическим свойствам вопрос принципиально не ориентирован на получение ответа и допускает отказ от ответа или отрицательный ответ.

В процессе создания тестового задания нередко имеет место несогласованность формируемого разработчиком утверждения с тем контекстом, который используется другим автором некоторого высказывания. Чтобы такие ситуации не возникали, надо пользоваться правилом П5.

П5: при подстановке в тестовое задание определяющего термина, полученный контекст должен оставаться истинным.

Тестовыми заданиями с кругом являются и утверждения: "материя - объективная реальность", "информация - сведения об объектах и явлениях окружающей среды", поскольку за определение, в соответствии с П6, можно принять только то утверждение, в котором нет разъяснения некоторого термина с помощью этого же понятия.

П6: тестовое утверждение не должно содержать логического круга.

При конструировании тестового задания определяющий признак должен быть указан непосредственно, а смысл и значение определяющих понятий должны быть хорошо представимы. Правило логичного мышления П7 указывает основу оперирования терминами.

П7: основные термины программно-педагогического тестового задания должны быть явно и ясно определены.

Чтобы избегать появления в тестовом задании лишних определяющих признаков, можно пользоваться правилом П8 и П9.

П8: в тестовом утверждении определяющий признак должен быть необходимым и достаточным.

П9: в тестовом утверждении определяющий признак должен быть существенным относительно рассматриваемого содержания учебной дисциплины.

При конструировании тестового задания необходимо предварительно уточнить проблемную ситуацию, которую необходимо разрешить в процессе формирования тестового задания в истинное высказывание. Для этого можно пользоваться правилом П10.

П10: истинность тестового высказывания должна оцениваться исходя из выбранных гносеологических предпосылок, определяемых спецификой области знаний.

Прагматически некорректным тестовым утверждением является задание с таким значением энтропии, которое является недоступным для формирования истинного ответа образом или с образом. В этом случае необходимо свести исходное тестовое задание с высокой энтропией к тестовым утверждениям с более низким уровнем неопределенности.

П11: сведение (декомпозиция) исходного тестового задания к утверждениям оптимальной энтропии необходимо выполнить в случае прагматически некорректного тестового задания.

Форма выражения содержания проблемной ситуации характеризует максимальную близость энтропии тестового задания к уровню достижений испытуемых заданного класса обученности (например, отличников). Отсюда следует методологическое правило:

П12: формулировка тестового задания детерминруется специфическими условиями задачи, которая решается с помощью данного задания, при условии возможно большей его различающей способности.

При рассмотрении форм тестовых заданий, они классифицируются на четыре группы, некоторые из которых разбиваются на виды. Классификация построена по признаку действий, выполняемых испытуемым в процессе формирования истинных высказываний:

Открытая форма *требует сформулированного самим обучаемым корректного ответа.* Имеет вид неполного утверждения, в котором отсутствует один элемент. Обычно тестируемый подставляет число или слово (допустимо словосочетание, состоящее не более чем из двух слов).

Закрыва́тая форма *предполагает выбор испытуемым правильного ответа из предложенных.* Состоит из неполного утверждения с одной вакансией и множества элементов, один или несколько из которых являются истинными. Испытуемый выбирает правильный ответ из предложенного множества.

Установить правильное соответствие — это, значит, *выбрать из двух приведенных множеств объектов истинные пары.* Имеет вид двух групп

элементов. Испытуемый должен связать каждый элемент первой группы с одним или несколькими элементами из второй группы. Рекомендуется дополнить вторую группу несколькими однотипными элементами, не связанными с первой группой.

Установить правильную последовательность означает *определение порядка следования* предложенных объектов (символов, слов, формул, рисунков). Дано множество неупорядоченных элементов, необходимо установить порядок между ними.

3.2.2. Формализованное описание механизмов формирования тестового задания

Функционально, конструктор тестовых заданий естественным образом разделяется на компоненты:

- графический редактор;
- формирование схемы разбора правильности ответа;
- параметризация выполнения (разрешенные действия);
- генерация Flash.

Функции графического редактора используются для создания графического образа с интерактивными полями.

$G = \{G_i\}_{i=1..J}$ – множество неактивных графических образов (картинка, формула и т.п., все что делается в других инструментальных средствах, текст, линия, стрелка и т.п., непосредственные функциональные возможности рисования в конструкторе)

$V = \{V_j\}_{j=1..J}$ – множество полей ввода для заданий открытого типа;

$Z = \{Z_n\}_{n=1..N}$ – множество полей захвата;

$A = \{A_m\}_{m=1..M}$ – множество альтернативных вариантов ответов.

Формирование схемы разбора правильности ответа вводит понятия позиций, которые связаны с множествами Z и A . Обозначим – P_Z и P_A . Множество дей D .

$D_1 - \langle A_i \mapsto P_{Z_i} \rangle$ - действие по перетаскиванию альтернативы;

$D_2 - \langle A_i \rightarrow P_{A_i} \rangle$ - изменение положения;

$D_3 - \langle \Rightarrow V_{A_i} \rangle$ - ввод текста;

$D_4 - \langle OK_i \rangle$ - подтверждение;

Закрытый тип $G = \{G_i\}_{i=1}^I$, Z_1 – единственное поле захвата,

$$\Lambda = \{A_m\}_{m=1}^M, D = D_1 \vee \{D_1 \vee D_2 \vee D_4\}.$$

Акцептор - $m^* \in \{1..M\}$.

Открытый тип - $G = \{G_i\}_{i=1}^I$, $V = V_1$ - единственное поле ввода, $D = \{D_3, D_4\}$

Акцептор - $\langle \text{строка символов} \rangle$

Классическое соответствие - $G = \{G_i\}_{i=1}^I$, $Z = \{Z_n\}_{n=1}^N$, $A = \{A_m\}_{m=1}^M$, $N=M$. $D = D_1 \vee \{D_1 \vee D_2 \vee D_4\}$.

Акцептор – перестановка на $\{1..M\}$ – соответствие альтернатив полям захвата.

Соответствие - $G = \{G_i\}_{i=1}^I$, $Z = \{Z_n\}_{n=1}^N$, $A = \{A_m\}_{m=1}^M$, $M > N$. $D = D_1 \vee \{D_1 \vee D_2 \vee D_4\}$.

Акцептор – упорядоченная последовательность из $\{1..M\}$ длины N – соответствие альтернатив полям захвата.

На упорядочение – полное соответствие на «классическое соответствие».

На порядок – полное соответствие на «соответствие».

На кластеризацию - $G = \{G_i\}_{i=1}^I$, $Z = \{Z_n\}_{n=1}^N$, $A = \{A_m\}_{m=1}^M$, $M > N$. $D = D_1 \vee \{D_1 \vee D_2 \vee D_4\}$.

Акцептор – отношение на $Z \times A$ $\{1..M\}$ длины N – соответствие альтернатив полям захвата. Матрица отношения имеет блочный вид.

На упорядоченное соответствие (перетаскивать определенные альтернативы в определенное место либо вводить определенный текст в определенном поле, и все это в определенной последовательности)

$$G = \{G_i\}_{i=1}^I, V = \{V_j\}_{j=1}^J, Z = \{Z_n\}_{n=1}^N, A = \{A_m\}_{m=1}^M, M > N. D = D_1 \vee D_2 \vee D_3$$

Правильность действий формализуется на основе задания конечного автомата

В случае жесткой последовательности он представляет вырожденный случай, где дуги помечаются либо некоторым действием ($\langle A_i \rightarrow P_{zi} \rangle, \langle \Rightarrow V_{Ai} \rangle$) либо его отрицанием ($\neg \langle A_i \rightarrow P_{zi} \rangle, \neg \langle \Rightarrow V_{Ai} \rangle$).

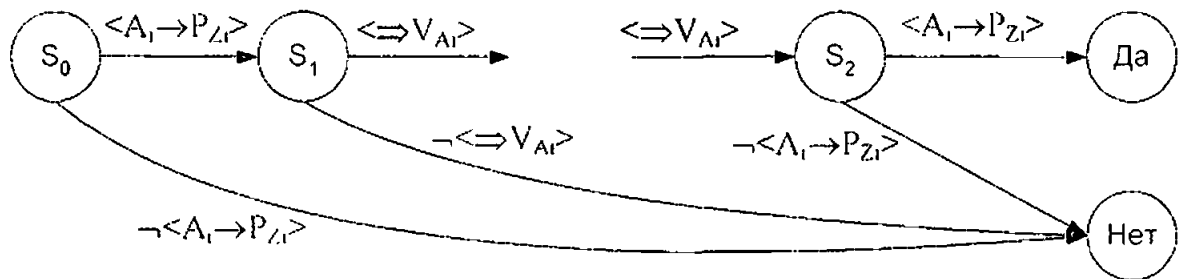


Рис. 3.3. Автоматная схема переходов

В общем случае, для реализации обобщающих функциональных возможностей необходимо формирование автомата произвольной структуры.

Детерминированный конечный автомат (ДКА) представляет собой пятерку - $\langle X, Y, \delta, y_0, F \rangle$, где X - конечный алфавит входных символов. Y - конечное множество состояний. $\delta: Y \rightarrow Y$ - функция переходов, $y_0 \in Y$ - начальное состояние. $F \in Y$ - множество допускающих состояний. Расширенная функция переходов $\hat{\delta}: X^* \times Y \rightarrow Y$, сопоставляющая новое состояние текущему состоянию и цепочке символов, определяется индуктивно следующим образом:

$$\forall y \in Y (\hat{\delta}(\varepsilon, y) = y), \forall y \in Y \forall \xi \in X^* \quad \forall x \in X (\hat{\delta}(\xi x, y) = \delta(x, \hat{\delta}(\xi, y))) \quad (3.7)$$

Если $\hat{\delta}(\xi, y_0) \in F$, то есть, стартуя в начальном состоянии и обработав цепочку ξ , автомат оказывается в одном из допускающих состояний. Ниже приведено автоматное описание работы проигрывателя тестов

Проигрыватель - $P = \langle S, E, A, T, S_b, S_f \rangle$

Состояния - $S = \{S_1, S_2, S_3, S_4, S_5, S_6\}$,

где: S_1 — останов; S_2 — проигрывание фрагмента; S_3 — пауза; S_4 — отображение списка фрагментов в режиме ручной паузы; S_5 — отображение списка фрагментов в режиме автоматической паузы; S_6 — отображение списка фрагментов в режиме останова.

Начальное состояние - $S_i = S_1$

Множество конечных состояний - $S_f = \{S_1\}$

События (входной алфавит): $E = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8, E_9, E_{10}\}$, где: E_1 — инициация начала или продолжения предъявления текущего фрагмента; E_2 — инициация приостановки предъявления текущего фрагмента; E_3 — инициация завершения предъявления текущего фрагмента; E_4 — инициация перехода к предыдущему фрагменту; E_5 — инициация перехода к следующему фрагменту; E_6 — инициация отката на один шаг назад по треку предъявления фрагментов; E_7 — инициация произвольного доступа к фрагменту; E_8 — наступление времени принудительного завершения предъявления текущего фрагмента; E_9 — инициация открытия окна со списком фрагментов; E_{10} — инициация закрытия окна со списком фрагментов.

Действия (выходной алфавит): $A = \{A_1, A_2, A_3, A_4, A_5, A_6, A_7\}$

A_1 — определение очередного фрагмента для проигрывания; A_2 — начало сбора статистики и запуск проигрывания фрагмента; A_3 — приостановка проигрывания фрагмента и приостановка сбора статистики; A_4 — возобновление проигрывания фрагмента и возобновление сбора статистики; A_5 — останов проигрывания фрагмента и завершение сбора статистики; A_6 — открытие списка фрагментов; A_7 — закрытие списка фрагментов.

Функция переходов $T(S_i, E_j) = \{ \langle (A_k)_{k=1..X}, S_m \rangle_{m=1..Y} \}$ ставит в соответствие каждой паре <состояние, событие> некоторую последовательность действий, которые проигрыватель выполняет при переходе в новое состояние, и само новое состояние (рис.3.4.).

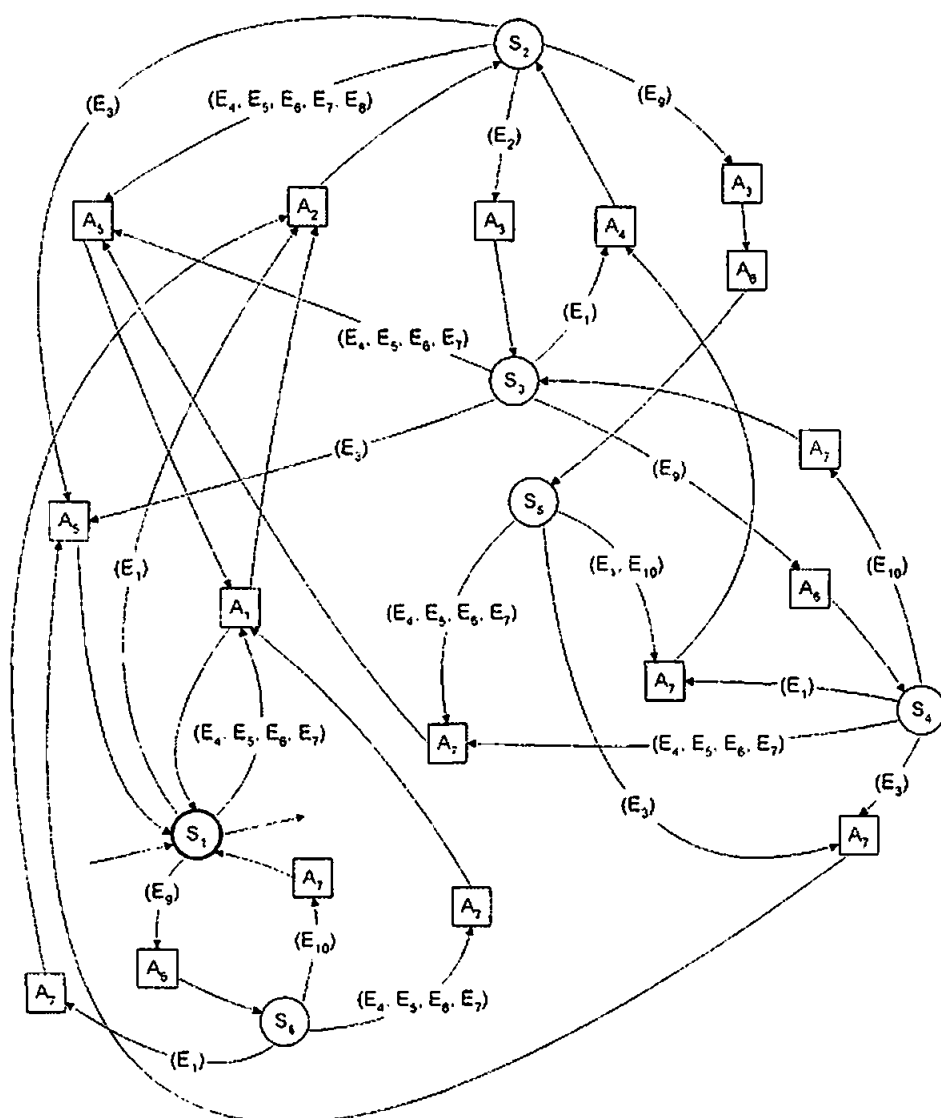


Рис. 3.4. Описание поведения проигрывателя тестовых заданий ДКА

Круг задач тестового контроля, охватывающий основные аспекты моделирования (для методиста с целью оценки его эффективности) и непосредственного контроля, решается путем создания совокупности сценариев, включающих:

- метод оценки вероятностей ошибочной классификации;
- ... (моделирующие алгоритмы)
- метод переоценки сложности тестовых заданий;
- язык формирования последовательности тестовых заданий;
- приложение компиляции языка;

- приложение разбора структуры гетерогенного теста;
- ...
- визуализация результатов тестового контроля;
- и другие.

3.3. Формализованная марковская модель поведения адаптивного алгоритма

Пусть адаптивный алгоритм представляет самый простой вариант, когда сложность задания увеличивается на единицу в случае верного ответа, и уменьшается на единицу в случае неверного. Если предположить, что сложность может быть задана некоторым числовым значением, то в результате возможна формализация процесса тестирования в виде марковской цепи, в которой вероятности переходов по сложностям определяются на основании логистической кривой.

3.3.1. Построение марковской цепи процесса адаптивного тестового контроля

Для формализованного представления модели тестирования будем использовать марковские цепи. Этот аппарат укладывается в общие концепции тестирования. Так, обычно предполагается, что ответы на задания - независимые величины. Поэтому можно использовать однородную марковскую цепь, где состояниями цепи являются меры сложности заданий.

Пусть $\xi = (C, P, \Pi)$,

где C - множество состояний марковской цепи (уровни сложности заданий);

P - матрица переходных вероятностей $P = \|p_{ij}\|$;

Π - вектор-строка начального распределения сложностей $\Pi = \|\pi_i\|$.

вероятности состояний C_i за π_i . При этих обозначениях вектор-строка стационарных вероятностей определяется равенством:

$$\bar{\pi} = (\pi_0, \pi_1, \pi_2, \dots, \pi_{I-1}, \pi_I). \quad (3.10)$$

Эти вероятности являются решением системы уравнений:

$$\bar{\pi} = \bar{\pi} \cdot P, \quad \sum_i \pi_i = 1. \quad (3.11)$$

Решение этой системы уравнений методом индукции даст результат:

$$\pi_i = \pi_0 \cdot \prod_{k=1}^{i-1} \frac{p_k}{q_{k+1}}, \quad i = \overline{1, I}, \quad \pi_0 = \left(1 + \sum_{j=0}^I \prod_{k=1}^{j-1} \frac{p_k}{q_{k+1}} \right)^{-1} \quad (3.12)$$

Используя уравнения переходных вероятностей (3.8) и (3.9), выраженные через логистическую функцию: $p_i = \Psi(\beta, \Theta)$; $q_i = 1 - \Psi(\beta, \Theta)$, $i = \overline{1, I}$, находим:

$$\begin{aligned} \pi_i &= \pi_0^i \prod_{k=1}^i \frac{\Psi(\Theta - \beta_k)}{1 - \Psi(\Theta - \beta_k)} = \pi_0^i \prod_{k=1}^i \frac{\exp(\Theta - \beta_k)}{1 + \exp(\Theta - \beta_k)} \left[1 - \frac{\exp(\Theta - \beta_k)}{1 + \exp(\Theta - \beta_k)} \right]^{-1} \\ &= \pi_0^i \prod_{k=1}^i \exp(\Theta - \beta_k) = \pi_0^i \exp \sum_{k=0}^i (\Theta - \beta_k) \end{aligned} \quad (3.13)$$

Т.е. в результате получим:

$$\pi_i = \pi_0 \exp \left(\sum_{k=1}^i (\Theta - \beta_k) \right), \quad i = \overline{1, I}. \quad (3.14)$$

Величины π_i^s являются распределением вероятностей на множестве состояний марковской цепи. Это стационарное распределение получено на основании введения некоторой дискретизации в пространстве «сложности». На рис.3.7. Приведены графики стационарных распределений для различного уровня знаний, когда уровень сложности соответствует номеру состояния.

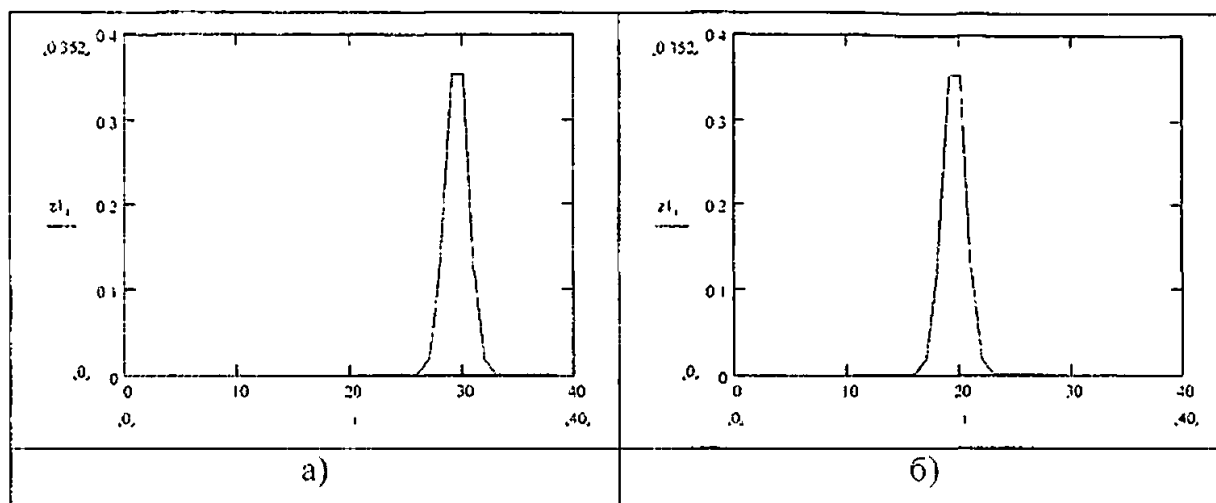


Рис.3.7. Стационарные распределения вероятностей.

Далее будем считать, что с каждым состоянием связана некоторая точка в пространстве «уровень знаний». В качестве меры разности между «уровнем» и «сложностью» будем использовать обозначение $z_i = \theta - \beta_i$. Здесь β_i - текущая сложность задачи. Считаем, что θ неизменно, но неизвестно, а β известно и управляемо.

3.3.2. Сходимость стационарных вероятностей при увеличении дискретизации шкалы уровня сложности

Поставим задачу увеличения дискретизации уровня сложности на той же шкале и предположим, что существует некоторая гипотетическая функция уровня знаний, положение которой мы должны найти.

ΔX - приращение сложности (в бальной шкале);

∇Y - градиент «знаний» в некотором универсальном множестве, в качестве оценки которого используется приращение функционала Y в локальной окрестности (ΔY).

Как было показано выше, управляемая модель задает дискретное распределение вероятностей на пространстве управляемых параметров. Будем предполагать, что смещенностью оценки целевой функции на k циклах можно пренебречь, тогда:

$$\forall i \quad M\hat{Y}(X_i) = Y(X_i) \text{ дисперсия } \sigma_k^2(X_i) = \frac{\sigma_1^2(X_i)}{k}$$

На основании этих предположений элементы матрицы переходных вероятностей марковской цепи **Ошибка! Источник ссылки не найден.**

можно переписать в виде

$$p_i = \Phi(Z_i) \quad q_i = 1 - \Phi(Z_i) \quad i = \overline{1, N} \quad Z_i = \frac{Y(X_i) - Y(X_{i-1})}{\sqrt{\sigma^2(X_i) + \sigma^2(X_{i-1})}}$$

Используя приближение нормального закона распределения логистической кривой получим:

$$\pi_i^X = \pi_0^X \prod_{k=1}^i \frac{\Phi^B(Z_k)}{1 - \Phi^B(Z_k)} = \pi_0^X \cdot e^{\beta \sum_{k=1}^i Z_k} \quad (3.15)$$

Величины π_i^X можно интерпретировать как распределение вероятностей на множестве $XX^C/X_0 = \{X_i\}_{i=1}^I$. Это множество дискретное. Далее будем исследовать предел распределения (3.15) при неограниченном увеличении дискретизации I ($I \rightarrow \infty$). Соответственно приращение параметра $\Delta_X \rightarrow 0$.

$$\forall i \geq 1 \quad \Delta_X = X_i - X_{i-1}. \quad (3.16)$$

Обозначим χ_i случайную величину на множестве $\{X_i\}_{i=1}^I$.

$$X_i = X_0 + i\Delta_X; \Delta_X = (X_M - X_0)/I, X_0 = \inf XX. \quad (3.17)$$

$X_M = \sup XX$ с распределением:

$$P\{\chi_i = X_i\} = \pi_i^X; \quad i = \overline{1, I}. \quad (3.18)$$

в котором π_0^X вычисляется из условной нормировки $\sum_{i=0}^I \pi_i^X = 1$.

Рассмотрим последовательность случайных величин χ_i , $I > 1$. χ_i имеет дискретное распределение. Если эта последовательность имеет предел, обозначим его χ . Случайная величина χ имеет непрерывное распределение на пространстве XX .

Утверждение 3.1. Последовательность случайных величин χ_1 сходится по функции распределения к χ с плотностью:

$$\omega(x) = \frac{1}{\alpha} \exp \left(\frac{\beta}{\sqrt{2}} \int_{x_0}^x \frac{Y'(u)}{\sigma(u)} du \right), \quad (3.19)$$

где α - нормирующий множитель.

Доказательство. Последовательность χ_1 сходится по распределению к χ [86], если для любой ограниченной непрерывной функции $f(X)$, $X \in XX$

$$\lim_{l \rightarrow \infty} Mf(\chi_l) = Mf(\chi). \quad (3.20)$$

Так как χ имеет непрерывное распределение, то:

$$Mf(\chi) = \frac{\int_{X_0}^{X_M} f(x) \exp \left(\frac{\beta}{\sqrt{2}} \int_{X_0}^x \frac{Y'(U)}{\sigma(U)} dU \right) dX}{\int_{X_0}^{X_M} \exp \left(\frac{\beta}{\sqrt{2}} \int_{X_0}^x \frac{Y'(U)}{\sigma(U)} dU \right) dX}. \quad (3.21)$$

Величины χ_1 дискретные, поэтому:

$$Mf(\chi_l) = \frac{\sum_{i=1}^l f(X_i) \cdot \exp \left(\beta \sum_{j=0}^i Z_j \right)}{1 + \sum_{i=1}^l \exp \left(\beta \sum_{j=0}^i Z_j \right)}. \quad (3.22)$$

Будем искать предел последнего выражения с учетом $\Delta X = \Delta X$, $\Delta X \rightarrow 0$. Рассмотрим предел суммы:

$$\lim_{\Delta X \rightarrow 0} \sum_{j=1}^l Z_j = \lim_{\Delta X \rightarrow 0} \sum_{j=1}^l \frac{Y'(x_j) + o(\Delta X)}{\sqrt{2\sigma^2(x_j) + o(\Delta X)}}. \quad (3.23)$$

Учитывая, что:

$$\frac{1}{\sqrt{2\sigma^2(x) + o(x)}} = \frac{1 + \sqrt{o(x)}}{\sqrt{2\sigma^2(x)}}. \quad (3.24)$$

Получим:

$$\lim_{\Delta \lambda \rightarrow 0} \sum_{j=1}^I z_j = \lim_{\Delta \lambda \rightarrow 0} \sum_{j=1}^I \frac{Y'(x_j)}{\sqrt{2} \cdot \sigma^2(x_j)} = \lim_{\Delta \lambda \rightarrow 0} \sum_{j=1}^I \frac{Y'(x_j) \Delta X}{\sqrt{2} \cdot \sigma(x_j)} = \frac{1}{\sqrt{2}} \int_{x_0}^x \frac{Y'(u)}{\sigma(u)} du. \quad (3.25)$$

Подставляя последнее выражение в (3.22) и домножая числитель и знаменатель на ΔX , получим:

$$\lim_{I \rightarrow \infty} Mf(\chi_I) = \lim_{I \rightarrow \infty} \frac{\sum_{i=1}^I f(X_i) \cdot \Delta X \cdot \exp\left(\frac{\beta}{\sqrt{2}} \int_{x_0}^x \frac{Y'(U)}{\sigma(U)} dU\right)}{\left(1 + \sum_{i=1}^I \exp\left(\frac{\beta}{\sqrt{2}} \int_{x_0}^x \frac{Y'(U)}{\sigma(U)} dU\right)\right) \cdot \Delta X}. \quad (3.26)$$

Учитывая интегрируемость функции $f(X)$ на XX , получим:

$$\lim_{I \rightarrow \infty} Mf(\chi_I) = \lim_{I \rightarrow \infty} \frac{\int_{x_0}^x f(X) \cdot \exp\left(\frac{\beta}{\sqrt{2}} \int_{x_0}^x \frac{Y'(U)}{\sigma(U)} dU\right) dX}{\int \exp\left(\frac{\beta}{\sqrt{2}} \int_{x_0}^x \frac{Y'(U)}{\sigma(U)} dU\right) dX}. \quad (3.27)$$

Таким образом, утверждение 3.1 доказано.

Таким образом, для дискретного распределения $\pi(X_i)$ может быть использовано приближение:

$$\omega^{\Pi}(x_i) = \frac{1}{\alpha} \exp\left(\frac{\beta}{\sqrt{2} \cdot \sigma} \int_{x_0}^x \frac{Y'(U)}{\sigma(U)} du\right) \quad (3.28)$$

где α - нормирующий множитель.

В итоге получено общее выражение для вычисления стационарного распределения значений управляемого параметра.

3.3.3. Плотность распределения сложности тестовых заданий

В результате доказанного утверждения для полученного дискретного стационарного распределения уровня сложности (3.14), в котором:

$$\pi_i = \pi_0 e^{\sum_{k=1}^i (\Theta - \beta_k)}, \quad i = \overline{1, I}. \quad (3.29)$$

При этом гипотетическая функция Y представляет интеграл $\Theta - \beta$.

$$Y(u) = \int_0^{\beta} (\Theta - u) du = \frac{(\Theta - u)^2}{2} \Big|_0^{\beta} = \frac{(\Theta - \beta)^2}{2}. \quad (3.30)$$

При этом распределение вероятностей принимает вид:

$$\pi(\beta) = \alpha \cdot \exp \frac{(\Theta - \beta)^2}{2}. \quad (3.31)$$

Таким образом, опять получили нормальное распределение, что подтверждает выбор нормального распределения для функции принадлежности уровня знаний.

3.3.4. Плотность распределения для функций регрессии с постоянной дисперсией и постоянным коэффициентом вариации

В практике моделирования наиболее часто оценки имеют либо постоянную дисперсию, либо постоянный коэффициент вариации. Поэтому рассмотрим частные случаи.

Постоянная дисперсия оценки

Предположим, что для всех значений управляемого параметра дисперсия оценки имеет вид $\forall X \in XX^c \sigma(X) = \sigma = Const.$

$$\omega^{\sigma}(x) = \frac{1}{\alpha(x_0)} \exp \frac{\beta \cdot Y(x)}{\sqrt{2} \cdot \sigma} \quad (3.32)$$

В этом случае выражение для плотности вероятности может быть представлено в следующем виде:

$$\omega^{\sigma}(x) = \frac{1}{\alpha} \exp \left(\frac{\beta}{\sqrt{2} \cdot \sigma} \int_{x_0}^x Y'(u) du \right) = \frac{1}{\alpha} \exp \left(\frac{\beta}{\sqrt{2} \cdot \sigma} [Y(x) - Y(x_0)] \right). \quad (3.33)$$

Вводя обозначение, независящее от X , перепишем формулу в виде:

$$\alpha(x_0) = \alpha \cdot \exp \left(-\frac{\beta}{\sqrt{2} \cdot \sigma} Y(x_0) \right). \quad (3.34)$$

Так как экспоненциальная функция монотонно возрастающая, то большим значениям целевой функции соответствуют большие значения

плотности распределения $W^\sigma(x)$. Соответственно максимум плотности приходится на максимум целевой функции.

В таблице 3.1. приведены значения плотности распределения $W^\sigma(x)$ при различных σ для целевой функции:

$$Y(X) = -0,36 \cdot (x - 5,0)^2 + 0,1 \quad (3.35)$$

Из графиков (рис.3.8.) видно, что чем меньше дисперсия, тем более выражено значение максимума. Так, например, для $\sigma=0.1$ модель будет пребывать 90% всего времени в интервале (3.0, 7.0), содержащем значение максимума.

Таблица 3.1.

X	W(x)/W(x*)				Y(x)
	$\sigma=0.1$	$\sigma=0.25$	$\sigma=0.5$	$\sigma=1.0$	
5.0	1.000	1.000	1.000	1.000	1.000
6.0	0.666	0.850	0.922	0.960	0.964
7.0	0.197	0.522	0.725	0.850	0.856
8.0	0.020	0.232	0.481	0.696	0.676
9.0	0.002	0.074	0.273	0.522	0.424
10.0	4·10 ⁻⁵	0.017	0.131	0.362	0.100

Таблица 3.2.

X	W(x)/W(x*)				Y(x)
	$\sigma=0.1$	$\sigma=0.25$	$\sigma=0.5$	$\sigma=1.0$	
5.0	1.000	1.000	1.000	1.000	1.000
6.0	0.666	0.850	0.922	0.960	0.964
7.0	0.197	0.522	0.725	0.850	0.856
8.0	0.020	0.232	0.481	0.696	0.676
9.0	0.002	0.074	0.273	0.522	0.424
10.0	4·10 ⁻⁵	0.017	0.131	0.362	0.100

Оценка с постоянным коэффициентом вариации

Предположим, что для всех значений:

$$\gamma = \frac{\sigma(x)}{Y(x)} = Const. \quad (3.36)$$

Значение интеграла в формуле (3.28) для плотности вероятности можно привести к виду:

$$\int_{x_0}^x \frac{Y'(u)}{\sigma(u)} du = \int_{x_0}^x \frac{dY(u)}{\gamma \cdot Y(u)} = \frac{1}{\gamma} [\ln|Y(u)| - \ln|Y(x_0)|]. \quad (3.37)$$

Отсюда получим следующее выражение:

$$\omega(x) = \frac{1}{\alpha} \exp \frac{1}{\gamma} \left[\ln \left| \frac{Y(x)}{Y(x_0)} \right| \right] = \frac{1}{\alpha} [Y(x)]^{\frac{1}{\gamma}} \quad (3.38)$$

В данном случае целевая функция может быть только положительной величиной. При $\gamma=1$ функция плотности вероятности с точностью до постоянного множителя на всем пространстве X совпадает с математическим ожиданием оценки функционала. Чем меньше коэффициент вариации, тем более выражен максимум плотности вероятности. В пределе при $\gamma=0$ распределение вероятности становится вырожденным и определяет сходимость алгоритмов поиска максимума.

В таблице 3.2. приведены значения $Y(x)$ и $W(x)$ на том же интервале и для той же функции (3.35), что и при постоянной дисперсии, только для различных значений коэффициентов вариации.

На рис. 3.9. приведены графики полученных зависимостей. Они показывают, что максимум плотности в случае постоянного коэффициента вариации носит более выраженный характер, чем для случая постоянной дисперсии оценки.

Таким образом, получены выражения (3.32) и (3.38), описывающие поведение управляющей модели в одномерном случае.

Для многомерного пространства управляемых параметров в общем случае не удастся получить аналогичных выражений.

Однако для аддитивной функции вида:

$$Y(\bar{x}) = \sum_{n=0}^N Y_n(x_n), \quad (3.39)$$

где $\bar{X} = (x_1, \dots, x_N)$ - вектор - строка; N- размерность пространства XX, распределения по координатам попарно независимо друг от друга.

Частные производные:

$$Y(\bar{x}) \frac{\partial Y}{\partial X_{\Pi}} = \frac{dY_{\Pi}}{dX_n} = Y'_n(X_n). \quad (3.40)$$

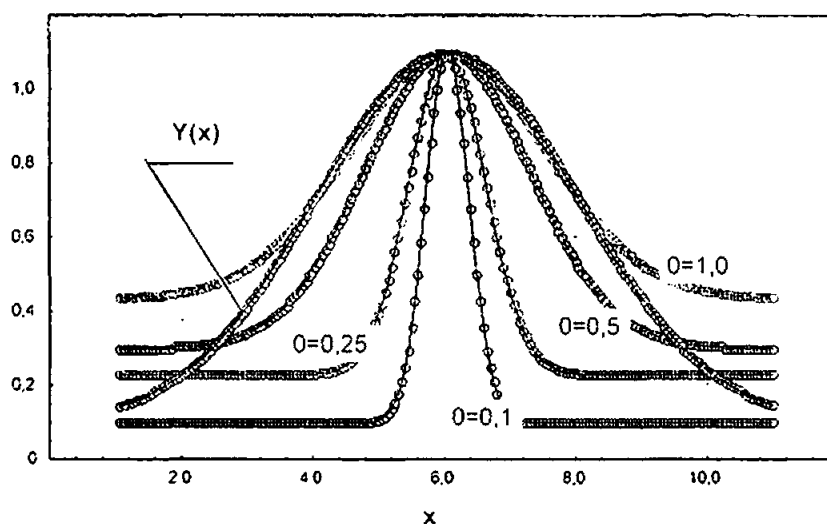


Рис.3.8. Постоянная дисперсия

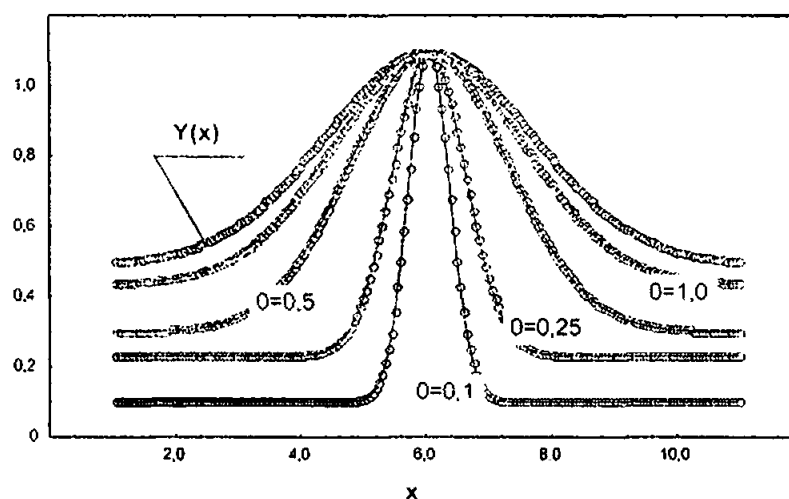


Рис.3.9. Постоянная вариация

В одномерном случае при вычислении плотности (3.32) учитывается лишь производная целевой функции, поэтому в случае постоянной дисперсии частные одномерные плотности определяются соотношениями:

$$\omega^{\sigma}(x) = \frac{1}{\alpha_n} \exp\left(-\frac{\beta}{\sqrt{2}} \frac{Y_n(x_n)}{\sigma}\right). \quad (3.41)$$

Таблица 3.3.

Вероятности ошибочной классификации

X _i	N(x _i)/N(x [*])				Y(x _i)
	γ=0.1	γ=0.25	γ=0.5	γ=1.0	
0.0	0.001	0.06	0.20	0.31	0.100
1.0	0.003	0.25	0.48	0.51	0.424
2.0	0.04	0.51	0.70	0.75	0.676
3.0	0.28	0.73	0.97	0.96	0.856
4.0	0.75	0.86	1.00	0.99	0.964
5.0	1.00	1.00	0.98	1.00	1.000
6.0	0.74	1.00	0.87	0.98	0.964
7.0	0.28	0.77	0.67	0.92	0.856
8.0	0.04	0.45	0.50	0.74	0.676
9.0	0.005	0.14	0.37	0.53	0.424
10.0	0.002	0.02	0.15	0.33	0.100

В силу независимости распределений общая плотность равна:

$$W(x) = \prod_{n=1}^N W_n(x_n) = \prod_{n=1}^N \frac{1}{\alpha_n} \exp\left(-\frac{\beta Y(x_n)}{\sqrt{2}\sigma}\right). \quad (3.42)$$

Вводя коэффициент:

$$\alpha^* = \prod_{n=1}^N \alpha_n, \quad (3.43)$$

и учитывая свойства экспоненциальной функции, получим:

$$\omega(\bar{x}) = \frac{1}{\alpha^*} \exp\left(-\frac{\beta}{\sqrt{2}} \frac{Y(\bar{x})}{\sigma}\right). \quad (3.44)$$

С целью проверки достоверности полученных результатов проводился имитационный эксперимент. Для функции (3.35) получены стационарные

распределения при различных дисперсиях. Результаты эксперимента приведены в таблице 3.2. Сравнение со значениями, полученными на основании (3.32) и приведенными в таблице 3.3. подтвердило обоснованность полученных аналитических соотношений.

Таким образом, разработанный алгоритм приводит к тому, что более точно оцениваются значения целевой функции, близкие к оптимальным.

В случае постоянного коэффициента вариации функция «уровня знаний» может быть только положительной величиной.

Таким образом, если есть мера «сложности» задания, то определена и мера «уровня знаний» и она совпадает с плотностью распределения адаптивного алгоритма тестового контроля.

Выводы по главе 3

1. Разработаны правила и принципы конструирования создания тестовых заданий. Сформированы требования к инструментальной среде создания интерактивных тестовых заданий с единых позиций автоматной схемы разбора корректности выполнения.

2. Разработан метод оценки весов для уровня сложности тестовых заданий, входящих в статический тест на базе процедур дискриминантного анализа. Проведен анализ влияния коррелированности ответов на веса сложности и разделяющую константу при классификации.

3. Разработана имитационная модель оценки матрицы ошибочных классификаций, параметризация которой включает размерность шкалы сложности, планы теста и механизмов предъявления, количественный состав группы тестируемых.

4. Построена марковская цепь процесса адаптивного тестового контроля, состояниями которой являются сложности заданий, а переходные вероятности определяются на основе логистической функции. Получены аналитические соотношения стационарных вероятностей цепи и показана их взаимосвязь с функцией уровня знаний.

4. ПРОГРАММНО-МОДЕЛИРУЮЩИЙ КОМПЛЕКС И АПРОБАЦИЯ СИСТЕМЫ АДАПТИВНОГО ТЕСТОВОГО КОНТРОЛЯ

В главе разрабатываются принципы создания программного комплекса поддержки процедур формирования тестовых заданий и алгоритмов адаптивного тестового контроля в системе переподготовки персонала транспортных предприятий. Приведена структура программного комплекса, построена инфологическая модель базы данных.

4.1. Проектирование структуры программного комплекса и интерфейса баз данных

При разработке структуры программного комплекса основной задачей было создание открытой системы с адаптивным интерфейсом и возможностью организации информационных связей с другими инструментальными средствами поддержки системы обучения по комбинированным технологиям.

4.1.1. Функции пользователей системы тестового контроля

При реализации функций системы выделим классы пользователей системы: администратор (А), методист (М), консультант (Р) и обучаемый (О). Все эти категории пользователей отличаются по их отношению к учебным материалам (U), которые определяют совокупность неделимых единиц учебной информации, соответствующих структуре знаний специализации, и тестовым заданиям (Т), которые необходимы для контроля уровня знаний обучаемых. Все указанные категории пользователей работают в одном информационном пространстве. Кроме того, руководители предприятий должны иметь возможность оперативного просмотра результатов переподготовки, что приводит к задаче сопряжения с системой мониторинга.

Аттестация сотрудников, как один из инструментов управления

персоналом. позволяет вырабатывать управленческие решения и решать вопросы не только управления кадрами, но и стратегического управления, планирования и развития.

Аттестации сотрудников транспортных предприятий представляют собой целенаправленный интерактивный, асинхронный процесс взаимодействия субъектов и объектов обучения между собой и со средствами обучения. Общиепринятыми компонентами системы аттестации являются подсистемы. целей обучения. содержания обучения. методов обучения, средств обучения. организационных форм обучения. а также идентификационно-контрольная, учебно-материальная, финансово-экономическая. нормативно-правовая, маркетинговая подсистемы. Эффективность может быть достигнута лишь за счет вложенной в систему адаптивности, гибкости, модульности, параллельности, технологичности и др

4.1.2. Организация диалога в системе тестового контроля

Обеспечение рационального диалогового взаимодействия в адаптивной системе тестового контроля достигается на основе выделения множества параметров диалога, существенных с точки зрения. как процесса компьютерного тестирования. так и психологически оптимальных условий общения

Сфера применения диалога охватывает процесс разработки ППТЗ, ППТ и технологий компьютерного тестирования. Реально компьютерный тестовый диалог нацелен на получение либо классификации испытуемых по уровням их подготовки. либо, на оценку уровня достижений тестируемых из заданного класса. Определение сферы применения диалога означает планирование того. какие по трудности ППТЗ разрешается системе предъявлять испытуемым.

По функции различают основной и вспомогательный диалог. Основной диалог относится к взаимодействию испытуемого с тестирующей системой и

реализует требуемый результат тестового контроля. Вспомогательный диалог (мегадиалог) предназначен для понимания компьютерного текста вводимого разработчиком ППТЗ, модификации, выбора и накопления банка заданий.

Вспомогательный диалог инициируется тестирующей системой с целью уточнения сообщения разработчика ППТЗ и администратора системы – это так называемый мегадиалог. Для облегчения вспомогательного диалога используется выбор из меню. Чтобы обеспечить реципиентам разработку и ввод ППТЗ используются специальные символы, располагаемые в верхней части экрана (окна).

Различают два вида адаптации компьютерного диалога: адаптация к ситуации выполнения ППТЗ и адаптация к испытуемому.

Адаптация системы учебной ситуации означает, что сообщения, которые испытуемый вводит в компьютер, анализируется в контексте решения учебной задачи, причем каждое сообщение испытуемого соотносится заложенной в программу соответствующим ответом. На основе анализа качества ответов принимается решение о соответствующем тестирующем воздействии.

Адаптация диалога к испытуемому (индивидуализация) осуществляется при наличии в системе модели конкретного испытуемого, и компьютер выдает контролирующее воздействие с ее учетом.

Индивидуализация диалога может осуществляться на нескольких уровнях. Первый уровень – учет истории ответов данного объекта, т.е. накопление и анализ данных по успешности его деятельности при выполнении ППТЗ, статистика его ошибок и выдача последующих контролирующих воздействий с учетом этой статистики. Модель испытуемого, реализованная в такой программе, весьма примитивна, поскольку в ней представлен только один чисто результативный параметр его учебной деятельности.

Следующий уровень предполагает обогащение модели испытуемого с учетом множества параметров его деятельности (стиль познавательной деятельности и т.п.)

Важнейшим параметром диалога, реализуемого тестирующей системой, является уровень его педагогической направленности. Такая направленность наблюдается в ситуациях, когда тест ориентирован на предметную экранную область, т.е. его цели не выходят за пределы выполнения теста и считается достигнутым, если получен конкретный результат – исходный балл испытуемого. При этом в случае затруднений программа осуществляет переключение модели объекта согласно имеющейся в семантической сети фреймам ситуаций (классам). В этом случае еще нет личностного оценивания, а есть управление структурой испытуемого.

Основа процесса компьютерного тестирования – оценка истинного балла конкретного испытуемого выступает здесь как побочный продукт деятельности объекта. Этот вид диалога, который применяется во многих системах компьютерного контроля часто без должного обоснования механически переносится на компьютерную технологию тестового контроля.

Рассматриваемый уровень педагогической направленности может быть реализован при классификации испытуемых на персональных компьютерах, в ЛВС и с применением интернет технологий. Его можно отнести к области самотестирования, результатом которого является отнесение конкретного испытуемого к определенному классу.

На втором уровне педагогической направленности диалога компьютерного тестового контроля его цели требуют выполнения ППТЗ, уровень трудности которых строго соотнесен к уровню способностей конкретного испытуемого.

Результатом компьютерного тестирования на втором уровне является индивидуальная оценка – истинный балл испытуемого данного класса обученности. В интеллектуальных тестирующих системах программа реализует определенную модель объекта, потом по мере диалога, эта модель

развивается, обучается и индивидуализируется в ходе анализа действий конкретного объекта.

Тесты, обладающие личностной отнесенностью к уровню подготовки испытуемого, адекватнее понимаются и формируют определенные отношения к принципу тестирования, преодолению препятствий, способности объекта оценивать собственную мыслительную деятельность, применение стратегий и т.п.

Данный вид диалога реализуется на персональных компьютерах или ЛВС при обеспечении присутствия преподавателя или администратора сети. Проведение личного компьютерного тестирования с итоговым оцениванием уровня достижений испытуемого и заключение юридически защищенных договоров по сети интернет практически невозможно. Это связано с тем, что вся деятельность в электронной сети основана целиком на доверии, ибо никаких правовых механизмов поддержки операций в сети интернет в настоящее время не существует. Авторизовать субъекта электронной сети, удаленного от банка ППТЗ, нереально.

Параметр симметричности диалога относится к мере инициативы, доступной для объекта и квазиобъекта по общению. Симметричный диалог предполагает равенство партнеров в инициировании шагов диалога. В компьютерном диалоге реально симметрия недостижима, поэтому можно говорить о разной степени асимметрии и ее дидактической целесообразности при разработке, модификации ППТЗ (тестов) и компьютерных технологий тестового контроля. При жесткой детерминации деятельности испытуемого в программу контроля заложена асимметрия диалога в пользу компьютера. Испытуемый получает возможность корректировки ответов до их ввода, выход из диалога в режиме анонимного или социально-ориентированного общения, когда сочтет это необходимым.

4.2. Структура базы данных системы тестового контроля

База данных локального варианта конструктора тестовых заданий реализована на Access, а сетевой вариант на базе MS SQL Server. Основные таблицы базы локального варианта представлены на рис.4.1. и 4.2.

Имя поля	Тип данных	Описание
Key	Счетчик	Ключ
Name	Текстовый	Название тестового задания
BackColor	Текстовый	Задний фон
BackImage	Текстовый	Полный путь к фоновой картинке
Type	Числовой	Режим тестирования (тренировка, контроль)
Vid	Числовой	Тип задания (перетаскивание, заполнение)
Font	Текстовый	Шрифт по умолчанию
Color	Числовой	Цвет шрифта по умолчанию
Size	Числовой	Размер шрифта по умолчанию
Bold	Логический	Жирный по умолчанию
Italic	Логический	Курсив по умолчанию
Underline	Логический	Подчеркивание по умолчанию
StrikeOut	Логический	Зачеркнутое по умолчанию
Transparent	Логический	Прозрачность
Level	Числовой	Уровень сложности задания
BackImageName	Текстовый	Имя фоновой картинки
FileName	Текстовый	Имя файла
File	Текстовый	Полное имя файла

Рис. 4.1. Таблица описания тестового задания

Имя поля	Тип данных	Описание
Key	Счетчик	Ключ
Name	Текстовый	Имя для генератора
Test	Числовой	Привязка к тестовому заданию
Type	Числовой	Тип объекта (текст, картинка, флэш-ролик)
Top	Числовой	Верхнее положение объекта
Left	Числовой	Левое положение объекта
Width	Числовой	Длина объекта
Height	Числовой	Высота объекта
Font	Текстовый	Шрифт надписи
Color	Числовой	Цвет шрифта объекта
Size	Числовой	Размер шрифта объекта
Bold	Логический	Жирный шрифт объекта
Italic	Логический	Курсив шрифта объекта
Underline	Логический	Подчеркнутый шрифт объекта
StrikeOut	Логический	Зачеркнутый шрифт объекта
Text	Поле MEMO	Надпись
Link	Числовой	Полное имя вставляемого файла
FileName	Текстовый	Имя файла

Рис. 4.2. Таблица описания объектов тестового задания

Реализована связь таблиц базы данных тестовых заданий с таблицами базы данных учебных материалов, представленная на рис.4.3. Основные данные располагаются непосредственно в таблице дисциплин (Discipl), также информация о цикле получается из таблицы циклов (Cycle), путем указания системного ключа цикла (c_id).

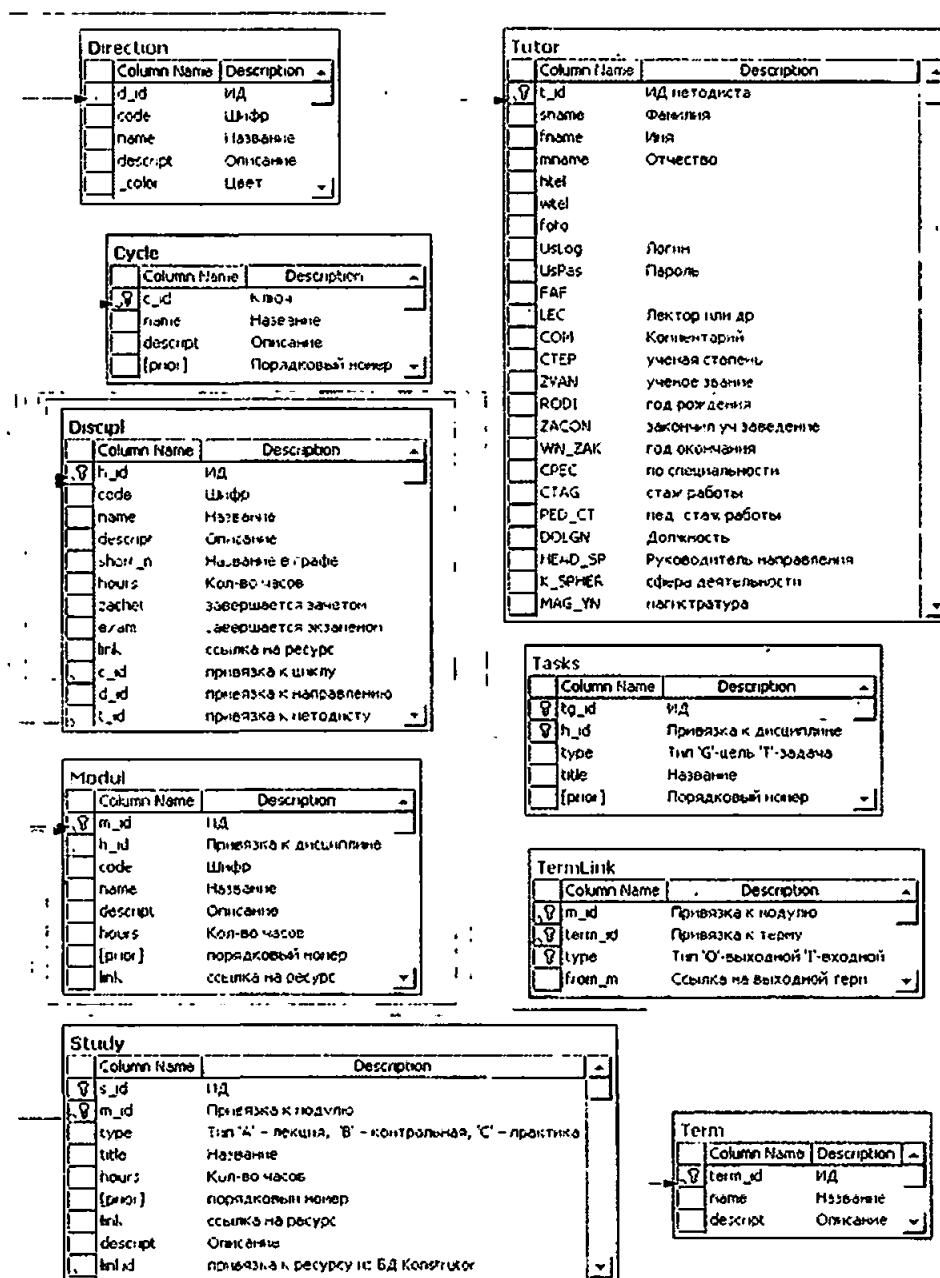


Рис. 4.3. Структура таблиц структуризации учебного материала

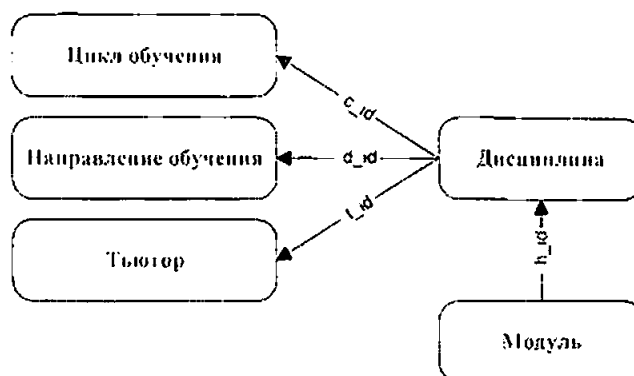


Рис. 4.4. Связь данных таблиц в описании дисциплины

Базовым элементом является таблица Модуль (Modul). Посредством ключевого поля (*m_id*), имеется возможность получить информацию о всех учебных элементах (лекциях, практических и тестовых заданиях). Структура информации о модуле имеет более сложный вид. Это связано с более сложной смысловой и методической нагрузкой на понятие модуль.

Доступ к термам (Term), осуществляется опосредовано, через таблицу-связку (TermLink), при этом если привязка входного термина производится на основе уже существующего выходного термина, то обязательно в поле (*from_m*) таблицы-связки указывается идентификатор модуля, где данный терм является выходным. Аналогично получению информации о терминах, осуществляется доступ к информации об учебных материалах (Resource) через таблицу-связку (Res_use).

Для работы системы в режиме карты требований к специализации создан набор таблиц, указанный на рис.4.5.

Данный набор данных предназначен для соотнесения требуемых знаний для получения определенной профессии (таблица Prof) и реальными знаниями, которые включаются в учебный план. Сравнение осуществляется на основе базовых понятий (термов) и количества часов.

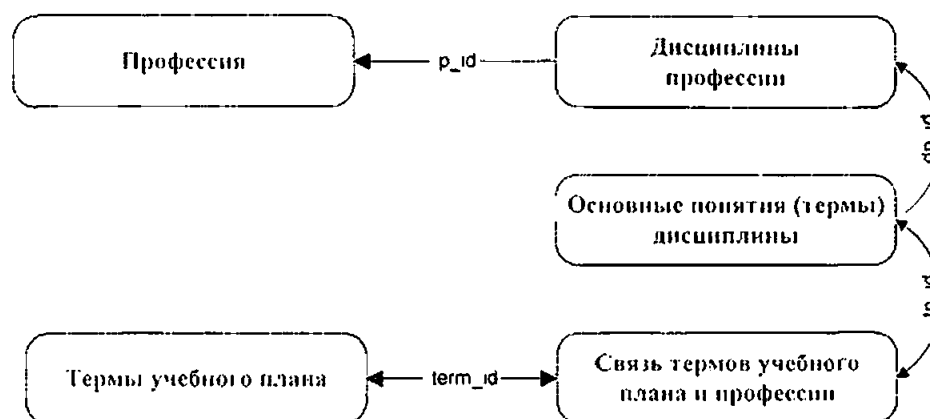


Рис. 4.5. Связь данных таблиц карты требований

Необходимость ведения функций контроля вызывает необходимость проработки системы защиты как информации о результатах аттестаций, так и учебно-методических материалов, которые являются собственностью предприятий. Разделение доступа в приложениях, управляющих базами данных, преследует две основные цели - недопущение (жесткое или мягкое) пользователя к тем функциям приложения, которые не входят в круг его функциональных обязанностей и предотвращение злоумышленного доступа к данным или их разрушения.

Права доступа к базе определяются на уровне сервера БД. Для облегчения работы администратора системы создан ряд таблиц и хранимых процедур, позволяющих автоматизировать процесс предоставления привилегий. Каждый преподаватель имеет свое имя пользователя и пароль. Каждый преподаватель имеет право изменять только свою дисциплину, либо, если он является руководителем направления, то и все остальные дисциплины этого направления. Все данные о преподавателях хранятся в таблице преподавателей. Единственным, кто имеет право изменять эту таблицу, является администратор. Администратор также имеет право изменять любую дисциплину. Также любой преподаватель, руководитель дисциплины либо администратор имеет право просматривать любую дисциплину, терм, модуль.

В диссертации реализованы механизмы интерфейсного взаимодействия с базовыми программными компонентами системы подготовки и структуризации учебных материалов, а именно, с конструктором лекций и конструктором курсов.

Конструктор лекций представляет инструментальную среду формирования лекций, как последовательности мультимедийных фрагментов, реализованных в различных инструментальных средах. Основная функция – формирование лекции, сбалансированной по форме представления учебного материала.

Конструктор курсов представляет инструментальную среду структуризации лекций в рамках единого учебного курса.

Среда Администратора позволяет регистрировать обучаемых, консультантов и методистов, изменять их личные данные и сведения результате прохождения аттестаций. Администратор также управляет взаимодействием обучаемых, консультантов, методистов между собой и их привязкой к учебному материалу.

Среда консультанта обеспечивает контроль процесса обучения каждого сотрудника. Позволяет проводить групповые и индивидуальные консультационные занятия и т.д.

Среда обучаемого обеспечивает все необходимые функции доступа к учебным материалам и взаимодействия с консультантами.

4.3. Анализ результатов аттестации по возрастным группам сотрудников

Анализ зависимости между уровнем системных профессиональных знаний работников и их возрастом основывается на данных по 307 работникам из 309, принявших участие в исследовании. Это связано с отсутствием по 2 работникам анкетных данных.

На основе данных о системных профессиональных знаниях работников, полученных в результате экзамена, было проведено исследование зависимости между уровнем знаний работников и их возрастом.

Анализ результатов экзамена по возрастным группам сотрудников проводился по двум направлениям:

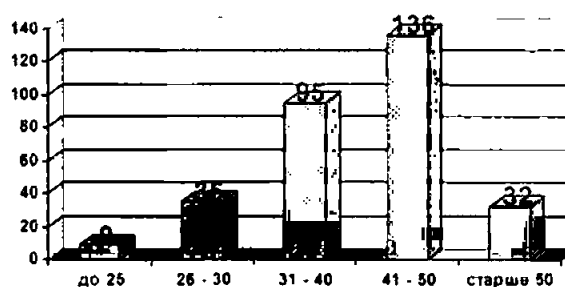
- анализ зависимости между уровнем системных профессиональных знаний и навыков финансово-экономического блока и возрастными группами сотрудников;
- анализ зависимости между результатами экзамена и возрастными категориями сотрудников по блокам: «Управление финансами», «Экономика», «Бухгалтерский учет и аудит», «Юридические основы финансово-экономического управления», «Организационные основы современного финансово-экономического управления» и «Информационные технологии».

4.3.1. Анализ зависимости по финансово-экономическому блоку

Для изучения зависимости между возрастом работников и их уровнем системных профессиональных знаний, работники были разделены на 5 возрастных групп: до 25 лет, 26-30 лет, 31-40 лет, 41-50 лет, старше 50 лет. Анализ распределения работников по возрастным группам (рис.4.6) показал, что наиболее многочисленными являются группы сотрудников в возрасте 41-50 лет (45%) и в возрасте 31-40 лет (31%). Самой малочисленной является группа сотрудников в возрасте до 25 лет (3%).

Для определения зависимости между уровнем системных профессиональных знаний работников и их возрастом был проведен корреляционный анализ.

Количественное соотношение
работников по возрастным группам



Процентное соотношение работников
по возрастным группам

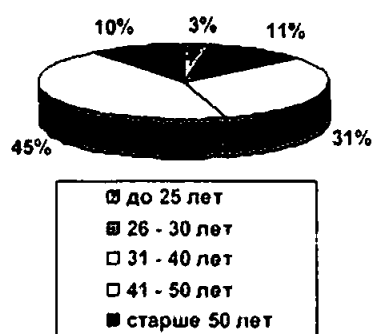


Рис. 4.6 Распределение работников по
возрастным группам

Таблица 4.1.

Значения коэффициента корреляции между возрастом работников и
их уровнем системных профессиональных знаний

Название блока	Значение коэффициента корреляции
Управление финансами	-0,23
Экономика	-0,25
Бухгалтерский учет и аудит	-0,05
Юридические основы финансово-экономического управления	-0,17
Организационные основы современного финансово-экономического управления	-0,21
Информационные технологии	-0,18
По финансово-экономическому блоку в целом	-0,28

Результаты корреляционного анализа представленные в табл.4.1. свидетельствуют о наличии слабой обратной связи между возрастом работника и его уровнем системных профессиональных знаний, которые были определены по результатам экзамена. как по отдельным блокам, так и по финансово-экономическому блоку в целом. Наибольшее значение коэффициента корреляции по модулю наблюдается по финансово-экономическому блоку в целом (-0.28), что тем не менее свидетельствует о наличии очень слабой обратной зависимости между возрастом работников и их уровнем системных профессиональных знаний. Следует отметить, что в рамках отдельных блоков знаний наибольшие по модулю значения коэффициента корреляции наблюдаются по блокам «Управление финансами» и «Экономика», наименьшее значение – по блоку «Бухгалтерский учет и аудит», что говорит практически о полном отсутствии связи между возрастом работников и их уровнем системных профессиональных знаний.

В результате обработки данных экзамена по финансово-экономическому блоку были определены уровни знаний и навыков по каждой возрастной группе.

В возрастной категории сотрудников до 25 лет максимальный балл по результатам экзамена составляет - 65,5, минимальный –33,6. Таким образом, разница между максимальной и минимальной оценкой результатов экзамена составляет 31,9 баллов.

В данной возрастной группе по результатам экзамена сотрудники имеют только два вида уровней знаний и навыков – средний и высокий. другие виды отсутствуют (рис 4.7).

По результатам экзамена 78% сотрудников показали высокий уровень знаний и навыков и 22% показали средний уровень знаний и навыков.

Достаточно высокий уровень системных профессиональных знаний по результатам экзамена в возрастной группе до 25 лет может быть связан с тем, что эта категория сотрудников недавно закончила обучение в средних и

высших учебных заведениях, в которых они имели возможность получить действительно современные и системные знания.

В возрастной категории сотрудников от 26 до 30 лет максимальный балл по результатам экзамена составляет 74,6, минимальный - 25,2. Таким образом, по сравнению с возрастной категорией сотрудников до 25 лет, эта группа имеет большее значение как максимального, так и минимального баллов, что увеличивает разницу между минимальной и максимальной оценками результатов экзамена, которая составляет 49,4 балла.

Распределение работников по уровню знаний и навыков в возрастной категории до 25 лет (включительно)



Процентное соотношение работников по уровню знаний и навыков в возрастной категории до 25 лет (включительно)

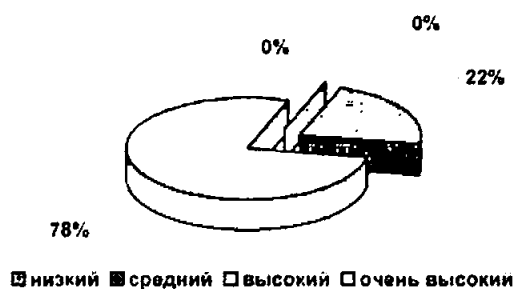


Рис. 4.7 Распределение работников по уровню знаний и навыков в возрастной категории до 25 лет

На рис.4.8 показано распределение работников по уровню знаний и навыков в возрастной категории от 26 до 30 лет. Также как и в возрастной

группе до 25 лет, сотрудники возрастной категории от 26 до 30 лет показали по результатам экзамена два вида оценок, которые соответствуют среднему и высокому уровню знаний и навыков.

Распределение работников по уровню знаний и навыков в возрастной категории от 26 до 30 лет (включительно)



Процентное соотношение работников по уровню знаний и навыков в возрастной категории от 26 до 30 лет (включительно)



Рис. 4.8 Распределение работников по уровню знаний и навыков в возрастной категории от 26 до 30 лет

Необходимо отметить, что если для категории сотрудников в возрасте до 25 лет преобладала доля сотрудников, имеющих высокий уровень знаний и навыков, то в возрастной группе от 26 до 30 лет доли сотрудников, имеющих средний и высокий уровни знаний практически равны. Увеличение количества сотрудников, имеющих средний уровень знаний и навыков, может быть связано с тем, что часть знаний, полученных при обучении в средних и высших учебных заведениях, могла устареть при отсутствии

повышения квалификации в этот период, а также на снижение результатов оценок могло повлиять снижение навыка сдачи экзаменов.

Повышение доли сотрудников, показавших средний уровень знаний и навыков

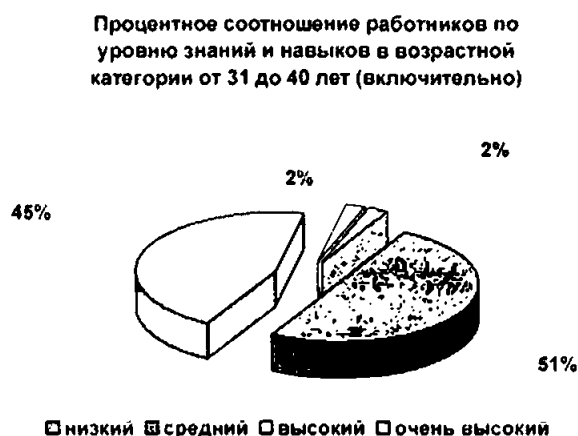
относительно возрастной группы до 25 лет может быть обусловлено тем, что сотрудники в этом возрасте уже имеют определенный практический опыт, который развивает и повышает практическую составляющую знаний и навыков.

В возрастной категории сотрудников от 31 до 40 лет максимальный балл по результатам экзамена составляет 76,1, минимальный - 20,9. Таким образом, разница между максимальным и минимальным значением результатов экзамена составляет 55,2 балла.

Как видно из рис.4.9 в возрастной категории от 31 до 40 лет сотрудники по результатам экзамена имеют все четыре уровня знаний и навыков. При этом сохраняется тенденция преобладания среднего и высокого уровня знаний (96% сотрудников данной возрастной категории). Однако необходимо отметить, что в отличие от двух предыдущих возрастных групп процент сотрудников, имеющих средний уровень знаний (51%) незначительно превышает процент сотрудников, имеющих высокий уровень знаний (45%). Количество сотрудников, имеющих очень высокий уровень знаний и навыков, незначительно и составляет только 2%.

На основе проведенного анализа результатов экзамена, можно сделать вывод о том, что для возрастной категории сотрудников от 31 до 40 лет характерно увеличение разницы между максимальными и минимальными результатами экзамена, по сравнению с предыдущими группами. Это связано с тем, что в данной возрастной группе появляются низкий и очень высокий уровни знаний. Наличие всех видов уровней знаний и навыков в данной возрастной группе может быть обусловлено тем, что с течением времени знания устаревают, поэтому при отсутствии своевременного их обновления они теряют свою актуальность (в связи с этим появляются сотрудники с

низким уровнем знаний и навыков и увеличивается доля сотрудников, имеющих средний уровень знаний и навыков). С другой стороны с возрастом у сотрудников, как правило, накапливается определенный опыт, который может дополняться обучением (в связи с этим появляются сотрудники с очень высоким уровнем знаний и увеличивается разрыв между максимальным и минимальным значением, полученных баллов).



Распределение работников по уровню знаний и навыков в возрастной категории от 31 до 40 лет (включительно)



Рис. 4.9 Распределение работников по уровню знаний и навыков в возрастной категории от 31 до 40 лет

В возрастной категории сотрудников от 41 до 50 лет максимальный балл по результатам экзамена составляет 71,5, минимальный - 17,3. Таким образом, разброс значений результатов экзамена составляет 53.2 балла.

Как видно из рис.4.10 в возрастной категории от 41 до 50 лет сотрудники по результатам экзамена имеют три уровня знаний и навыков. При этом сохраняется преобладание группы сотрудников (96%) , имеющих средний и высокий уровни знаний и навыков.



Рис. 4.10 Распределение работников по уровню знаний и навыков в возрастной категории от 41 до 50 лет

Необходимо отметить, что сохраняется тенденция увеличения доли сотрудников со средним уровнем знаний (58%, что на 7% больше, чем в предыдущей группе) и низким уровнем знаний (6%, что на 4% больше, чем в предыдущей группе) при переходе к более высокой возрастной категории сотрудников. Вместе с тем, в данной возрастной группе произошло увеличение разницы между количественным соотношением групп

сотрудников со средним и высоким уровнем знаний и навыков , который составил 20%.

Некоторое увеличение доли сотрудников с низким и средним уровнем знаний может быть связано с устареванием знаний, которое появляется в случае отсутствия налаженной системы обучения персонала.

В возрастной категории сотрудников свыше 50 лет максимальный балл по результатам экзамена составляет 66,7, минимальный -14,4. Таким образом, разница между максимальным и минимальным значением результатов экзамена составляет 52,4 балла.

Как видно из рис.4.11 в возрастной категории свыше 50 лет сотрудники по результатам экзамена имеют три уровня знаний и навыков. При этом сохраняется преобладание группы сотрудников (87%), имеющих средний и высокий уровни знаний и навыков. Необходимо отметить, что сохраняется тенденция увеличения доли сотрудников со средним уровнем знаний (65%, что на 7% больше, чем в предыдущей группе) и низким уровнем знаний (13%, что на 9% больше, чем в предыдущей группе) при переходе к более высокой возрастной категории сотрудников. Вместе с тем, в данной возрастной группе произошло дальнейшее увеличение разницы между количественным соотношением групп сотрудников со средним и высоким уровнем знаний и навыков, которая составил 43%. Разница между группой сотрудников имеющих низкий уровень знаний и навыков и группой сотрудников , имеющих высокий уровень знаний и навыков составила всего лишь 9%. В данной возрастной категории отсутствуют сотрудники, которые имеют очень высокий уровень знаний и навыков.



Рис. 4.11 Распределение работников по уровню знаний и навыков в возрастной категории свыше 50 лет

Еще большее увеличение числа сотрудников, соответствующих среднему и низкому уровню знаний и навыков в возрастной категории свыше 50 лет, может быть обусловлено тем, что система повышения квалификации персонала на предприятии либо не охватывает эту возрастную категорию сотрудников, либо на предприятии система повышения квалификации персонала недостаточно увязана с системой стимулирования сотрудников, либо сотрудникам в текущей деятельности не требуются знания в смежных областях.

4.3.2. Анализ зависимости между результатами аттестации и возрастными категориями сотрудников по блокам

Распределение набранных баллов по блокам по каждой возрастной категории представлено на рис.4.12. Анализ полученных результатов показал, что категория сотрудников до 25 лет имеет более высокий уровень знаний и навыков относительно других групп по всем блокам, кроме блоков «Бухгалтерский учет и аудит» и «Экономика».

Наиболее низкие результаты были показаны всеми возрастными группами по блоку «Бухгалтерский учет и аудит». Полученные баллы соответствуют среднему уровню знаний и навыков. Категория сотрудников старше 50 лет показала более низкий уровень знаний и навыков относительно других групп по всем блокам, кроме блока «Бухгалтерский учет и аудит».

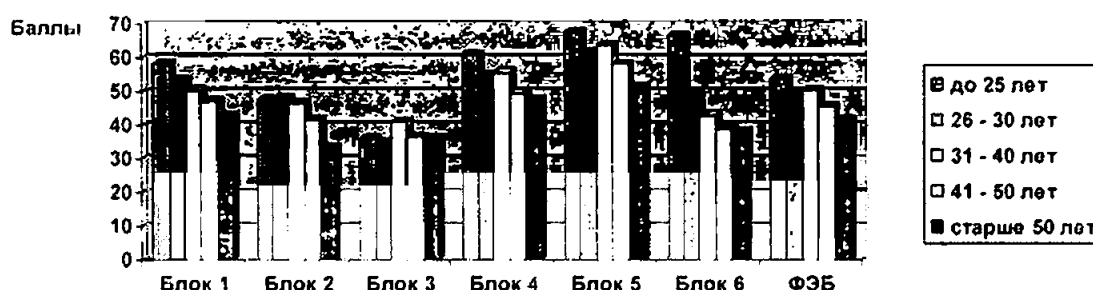


Рис. 4.12 Распределение набранных баллов по блокам и возрастным группам

Такое распределение результатов может быть связано с тем, что, несмотря на постоянные нововведения в области бухгалтерского учета и аудита, уровень знаний и навыков по этому направлению повышается по мере накопления практического опыта, а это происходит с увеличением возраста.

Наиболее явно проявляется преимущество возрастной категории до 25 лет по блоку «Информационные технологии». Средний балл по этому блоку

составил 66,6, что соответствует высокому уровню знаний и навыков. Остальные возрастные группы показали средний уровень знаний и навыков по этому блоку, при последовательном снижении средних баллов по мере перехода от группы к группе

(с 48,7 баллов для возрастной категории от 25 до 30 лет и до 38,4 баллов для возрастной категории старше 50 лет). Полученные результаты могут быть обусловлены тем, что сотрудники наиболее молодой возрастной группы недавно получили необходимые знания и навыки по этому блоку во время обучения в средних и высших учебных заведениях, а сотрудники остальных возрастных групп либо не имели возможности получить необходимые знания и навыки, либо они не были востребованы и в настоящее время забыты.



Рис. 4.13 Распределение баллов сотрудников по блокам в возрастной категории до 25 лет

На рис.4.13 показано распределение баллов сотрудников по блокам в возрастной категории до 25 лет. Наиболее высокий результат категория сотрудников в возрасте до 25 лет показала в блоках «Юридические основы финансово-экономического управления» (61 балл, что соответствует высокому уровню знаний и навыков), «Организационные основы современного финансово-экономического управления» (67,9 баллов, что

соответствует высокому уровню знаний и навыков), «Информационные технологии» (66,6 баллов, что соответствует высокому уровню знаний и навыков).

Наиболее низкий результат эта возрастная группа показала по блоку «Бухгалтерский учет и аудит – 36,1 балла, что свидетельствует о недостаточной подготовке сотрудников по этому направлению профессиональной деятельности.

На рис.4.14 показано распределение баллов сотрудников по блокам в возрастной категории от 26 до 30 лет. Анализ полученных результатов позволил сделать следующие выводы:

- уровень знаний и навыков этой возрастной группы является сбалансированным по блокам;
- только по трем блокам («Управление финансами», «Юридические основы финансово-экономического управления», «Организационные основы современного финансово-экономического управления») данная возрастная категория имеет высокий уровень знаний и навыков, при этом необходимо отметить, что результаты приближаются к нижней границе этого уровня;
- по блокам «Экономика», «Бухгалтерский учет и аудит», и «Информационные технологии» данная возрастная категория имеет средний уровень знаний и навыков.

В целом можно сделать вывод о том, что данная возрастная категория несколько уступает по полученным результатам сотрудникам возрастной категории до 25 лет.



Рис. 4.14 Распределение баллов сотрудников по блокам в возрастной категории от 26 до 30 лет

На рис.4.15 показано распределение баллов сотрудников по блокам в возрастной категории от 31 до 40 лет. Анализ полученных результатов позволил сделать следующие выводы:

- уровень знаний и навыков этой возрастной группы является сбалансированным по блокам ;
- по четырем из шести блоков («Управление финансами», «Экономика», «Бухгалтерский учет и аудит», «Информационные технологии») полученные результаты соответствуют среднему уровню знаний и навыков;
- высокий уровень знаний и навыков данная возрастная группа имеет только по блокам «Юридические основы финансово-экономического управления» и «Организационные основы современного финансово-экономического управления»)

В целом можно сделать вывод о том, что данная возрастная категория несколько уступает по полученным результатам сотрудникам возрастной категории до 25 лет и соответствует уровню знаний и навыков сотрудников возрастной категории от 26 до 30 лет.



Рис. 4.15 Распределение баллов сотрудников по блокам в возрастной категории от 31 до 40 лет

На рис.4.16 показано распределение баллов сотрудников по блокам в возрастной категории 41 до 50 лет. Анализ полученных результатов позволил сделать следующие выводы:

- уровень знаний и навыков этой возрастной группы является сбалансированным по блокам ;
- только по одному блоку «Организационные основы современного финансово- экономического управления» эта возрастная категория сотрудников показала высокий уровень знаний и навыков (58,2 балла).

В целом можно сделать вывод о том, что данная возрастная категория имеет средний уровень знаний и навыков, который совпадает с возрастной категорией от 31 до 40 лет и уступает по полученным результатам сотрудникам возрастной категории до 25 лет.



Рис. 4.16 Распределение баллов сотрудников по блокам в возрастной категории от 41 до 50 лет

На рис.4.17 показано распределение баллов сотрудников по блокам в возрастной категории старше 50 лет. Анализ полученных результатов позволил сделать следующие выводы:

- уровень знаний и навыков этой возрастной группы является сбалансированным по блокам ;
- только по одному блоку «Организационные основы современного финансово-экономического управления» данная возрастная категория показала высокий уровень знаний и навыков. при этом полученные результаты приближаются к нижней границе оценки (51,8 балла при интервале допустимых значений для этого уровня 50,1-75 баллов);
- оценки по остальным пяти блокам соответствуют среднему уровню знаний и навыков.

Распределение работников по количеству набранных баллов

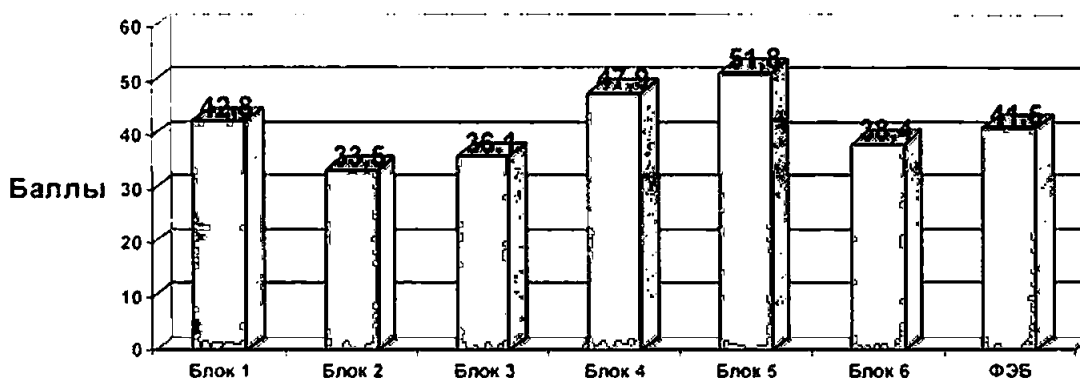


Рис. 4.17 Распределение баллов сотрудников по блокам в возрастной категории старше 50 лет

В целом можно сделать вывод о том, что данная возрастная категория имеет средний уровень знаний и навыков и уступает по полученным результатам всем остальным возрастным категориям сотрудников.

Выводы по главе 4

1. Разработана общая структура программного комплекса поддержки адаптивных механизмов предъявления тестовых заданий, включая вопросы проектирования базы данных и ограничений прав доступа с учетом выбранной классификацией пользователей системы.

2. Разработана инструментальная среда конструирования графических интерактивных тестовых заданий, позволяющая формировать задания «на соответствие», «упорядочение», «открытой» формы и «закрытой».

3. Проведены имитационные эксперименты и показана эффективность разработанных методов и алгоритмов.

ЗАКЛЮЧЕНИЕ

1. Проведен системный анализ процедур тестового контроля. Рассмотрены проблемы кадрового обеспечения промышленных предприятий и общие тенденции развития системы непрерывного образования. Рассмотрены математические модели и методы моделирования процессов аттестации персонала, которые позволяют повысить эффективность системы аттестации персонала промышленных предприятий.

2. Показано, что адаптивный механизм предъявления тестовых заданий обеспечивает выбор каждого последующего задания из базы тестовых заданий в зависимости от ответов на все предыдущие, что дает большую свободу построения новых методических принципов тестового контроля.

3. Для осуществления мониторинга уровня сформированности требуемых характеристик специалиста в диссертации предлагается использовать комплексный показатель качества психолого-индивидуальных компетенций включающий локальные коэффициенты сформированности профессиональных компетенций, что позволяет более дифференцированно подойти к оценке уровня знаний.

4. Рассмотрена схема группового тестирования с произвольным количеством участников различной подготовленности. При этом матрица ответов представляет собой ту исходную информацию, по которой предполагается оценить латентные параметры тестирования, а именно, трудности заданий и уровни подготовленности испытуемых одновременно, что повышает адекватность моделей тестового контроля.

5. Дается формализованное описание методики конструирования тестовых заданий, с использованием разработанного конструктора тестовых заданий, который имеет следующие функциональные компоненты: графический редактор формирования структуры задания; формирование схемы разбора ответа; параметризация выполнения задания; подсистема генерации Flash-файла тестового задания.

6. Предложенный в работе круг вопросов организации тестового контроля охватывает основные аспекты моделирования, оценки эффективности и непосредственного контроля, что решается путем создания совокупности сценариев, включающих: метод оценки вероятностей ошибочной классификации; метод переоценки сложности тестовых заданий; язык формирования последовательности тестовых заданий и другие приложения.

7. Сформированы требования к реализации программно-моделирующего комплекса тестового контроля и аттестации персонала предприятий промышленности и транспортного комплекса. Разработанный программный комплекс, методы и алгоритмы прошли апробацию и внедрены для практического применения в ряде предприятий, а также используются в учебном процессе на кафедре АСУ МАДИ.

Литература

1. Астанин С.В. Мониторинг процесса обучения в системе открытого образования // Интеллектуальные САПР. Таганрог, 2001. №4.
2. Астанин С.В., Захаревич В.Г., Попов Д.И. Интеллектуальные средства обучения в Интернет // Сборник докладов Всероссийской научной конференции «Управление и информационные технологии». СПб., 2003. Т. 2. С. 278–282.
3. Астанин С.В., Калашникова Т.Г. Разработка индивидуальной модели поведения обучаемого в системе дистанционного образования // Перспективные информационные технологии и интеллектуальные системы. Таганрог, 2001. №5.
4. Астанин С.В., Курейчик В.М., Попов Д.И., Кузьмицкий А.А. Интеллектуальная образовательная среда дистанционного обучения // Новости искусственного интеллекта. М., 2003. № 1.03 (55). С.7–14.
5. Баяндин Д.В., Кубышкин А.В., Мухин О.И., Рябуха А.А. Математическое моделирование в системе "Stratum Computer" // Труды Всероссийской научно-практической конференции "Проблемы образования, научно-технического развития и экономики Уральского региона". Березники, 1996, с.80-81.
6. Безкоровайный М.М., Костогрызов А.И., Львов В.М. Инструментально-моделирующий комплекс для оценки качества функционирования информационных систем «КОК». Руководство системного аналитика. – М.: Синтег, 2000. – 116с.
7. Бенькович Е.С., Колесов Ю.Б., Сениченков Ю.Б. Практическое моделирование сложных динамических систем. С. Петербург, БХВ, 2001.- 441с.
8. Бершадский А.М., Кревский И.Г. Дистанционное обучение – форма или метод // Дистанционное образование. М., 1998. № 4.
9. Беспалько В.П. Теория учебника: Дидактический аспект. М.: Педагогика, 1988.

10. Бизли Д.. Язык программирования PYTHON, Киев, ДиаСофт, 2000. – 336 с.
11. Богге У, Богге М. UML и Rational Rose, М.: Лори, 2000. - 582с.
12. Болотник Л.В., Соколова М.А. Тематическая модель структуры учебного материала // Проблемы педагогических измерений: Межвуз. сб. тр. / Под ред. В. И. Левина. М., 1984.
13. Буравлев А.И., Переверзев В.Ю. Выбор оптимальной длины педагогического теста и оценка надежности его результатов // Дистанционное образование. М., 1999. № 2. С. 27.
14. Буч Г. Объектно-ориентированный анализ и проектирование с примерами на C++, 3-е изд. / Пер. с англ. – М.: «Издательство Бином», СПб.: «Певский диалект», 2001 – 560с.
15. Буч Г., Рамбо Д., Джекобсон А. Язык UML. Руководство пользователя: Пер. с англ. – М.: ДМК, 2000. – 432с.
16. Васильев А.Е., Киричков А.В., Леонтьев А.Г. Исследование мехатронных систем с нечетким управлением с применением пакета Model Vision 3.0. // Гибридные системы. Model Vision Studium: Труды междунар. науч.-технич. конф. СПб.: Изд-во СПбГТУ, 2001. – с.46-50.
17. Васильев А.Е., Леонтьев А.Г. Применение пакета Model Vision Studium для исследования мехатронных систем. // Гибридные системы. Model Vision Studium: Труды междунар. науч.-технич. конф. СПб.: Изд-во СПбГТУ, 2001. – с.51-52.
18. Васильев В.И., Демидов А.Н., Малышев Н.Г., Тягунова Т.Н. Методологические правила конструирования компьютерных педагогических тестов. М.: Изд-во ВТУ, 2000.
19. Васильев В.И., Тягунова Т.Н. Основы культуры адаптивного тестирования. – М.: Издательство ИКАР, 2003. 584 с.
20. Вендров А.М. CASE-технологии: Современные методы и средства проектирования информационных систем. М.: Финансы и статистика, 1998. - 176с.

21. Вишняков Ю.М., Кодачигов В.И., Родзин С.И. Учебно-методическое пособие по курсам «Системы искусственного интеллекта», «Методы распознавания образов». Таганрог: Из-во ТРТУ, 1999.
22. Гаврилова Т.А. Червинская К.Р. Извлечение и структурирование знаний для экспертных систем. М.: Радио и связь, 1992. 200 с.
23. Георгиев В.О. Модели представления знаний предметных областей диалоговых систем // Изв. АН СССР. Техн. киберн. 1993. №5.
24. Голец И.Н., Попов Д.И. Модель представления знаний в интеллектуальной системе дистанционного образования // Известия ТРТУ. Тематический выпуск. Интеллектуальные САПР. Таганрог, 2001. С. 332 – 336.
25. Голец И.Н., Попов Д.И. Формальный способ представления знаний в дистанционном образовании // Материалы III Всероссийской научной конференции молодых ученых и аспирантов «Новые информационные технологии. Разработка и аспекты применения». Таганрог, 2000. С. 158 – 159.
26. Гома Х. UML. Проектирование систем реального времени, параллельных и распределенных приложений: Пер. с англ. – М.: ДМК Пресс, 2002. – 704с.
27. Гуленко В.В. Формы мышления. // Соционика, ментология и психология личности, N 4, 2002 (<http://socio16.narod.ru/v/gul-402.html>).
28. Гулытьев А.К.. MATLAB 5.3. Имитационное моделирование в среде Windows, М.: Корона принт, 2001. – 400с.
29. Дая У., Мюрхауг Б., Ньюгород К. СИМУЛА-67. Универсальный язык программирования. М.: Мир, 1969. - 99с.
30. Дмитриев А.К., Мальцев П.А. Основы теории построения и контроля сложных систем. Л.: Энергоатомиздат, 1988.- 192 с.
31. Дьяконов В. Mathematica 4: учебный курс. – СПб: Питер, 2002. – 656 с

32. Емельянов С.В., Коровин С.К. Новые типы обратной связи. М.: Наука, 1997. 352 с.

33. Зегжда Д.П., Иванко А.М. Основы безопасности информационных систем М.: Горячая линия – Гелеком, 2000. 452 с.

34. Калашникова Т.Г. Исследование и разработка методов и моделей правдоподобных рассуждений в интеллектуальных системах поддержки принятия решений. Диссертация на соискание ученой степени кандидата технических наук. Таганрог, 2001.

35. Карташев М.И. Методы и алгоритмы адаптивного компьютерного тестирования / Баринов К.А., Рогова О.Б., Карташев М.И. // В мире научных открытий № 9 (21). – Красноярск: НИЦ, 2011. – С. 93-106.

36. Карташев М.И. Адаптивный тестовый контроль в системах дистанционного образования / Никитин М.М, Строганов В.Ю., Карташев М.И. // В мире научных открытий № 9 (21). – Красноярск: НИЦ, 2011. – С. 118-126.

37. Карташев М.И. Организационно-технические принципы проектирования иерархических структур управления промышленными объединениями / Карасев А.А., Карташев М.И., Котов А.А., Луковецкая Т.М., Чугунова Д.Н. // Автоматизация управления в организационных системах: межвуз. сб. науч. тр. МАДИ (ГТУ). – М.: МАДИ (ГТУ), 2008. – С. 27-37.

38. Карташев М.И. Формирование организационной структуры в Business Studio / Карташев М.И., Приходько М.В., Свободин В.Ю., Строганов Д.В. // Автоматизация управления в организационных системах: межвуз. сб. науч. тр. МАДИ (ГТУ). – М.: МАДИ (ГТУ), 2008. – С. 58-68.

39. Карташев М.И. Использование нечетких множеств при определении количественных оценок связности учебного материала / Николаев А.Б., Ягудаев Г.Г., Карташев М.И., Ульянова А.И., Свободин В.Ю. // Интерактивные технологии моделирования и управления: сб. науч. тр. МАДИ № 2/46. – М.: МАДИ, 2010. – С. 128-134.

40. Карташев М.И. "Интеграция приложений в системе ""СОТА"" / Ягудаев Г.Г., Карташев М.И. // Методы описания и моделирования бизнес-процессов и технологий в промышленности, строительстве и образовании: сб. науч. тр. МАДИ № 3/47. – М.: МАДИ, 2010. – С. 97-104.

41. Карташев М.И. Интерфейсные формы системы управления персоналом / Брыль В.Н., Карташев М.И., Николаев А.Б., Свечников А.А., Якунин П.С. // Автоматизация систем управления персоналом: сб. науч. тр. МАДИ. – М.: МАДИ, 2011. – С. 11-16.

42. Карташев М.И. Проблемы принятия решений при формировании и оценивании эффективности бизнес-процессов / Брыль В.Н., Горячкин Б.С., Карташев М.И., Строганов Д.В., Якунин П.С. // Автоматизация систем поддержки управленческой деятельности: сб. науч. тр. МАДИ. – М.: МАДИ, 2011. – С. 21-28.

43. Карташев М.И. Семантическое моделирование данных в системах поддержки управленческой деятельностью / Карасев А.А., Карташев М.И., Тимофеев П.А., Ягудаев Г.Г. // Автоматизация систем поддержки управленческой деятельности: сб. науч. тр. МАДИ. – М.: МАДИ, 2011. – С. 77-81.

44. Карташев М.И. Методика расчета экономической эффективности инновационных проектов с учетом фактора риска / Батов Р.В., Васильев Д.А., Карташев М.И., Катырин С.Н., Ягудаев Г.Г. // Автоматизация и управление: стратегия, инвестиции, инновации. сб. науч. тр. МАДИ. – М.: Техполиграфцентр, 2011. – С. 36-41.

45. Карташев М.И. Оценка влияния начальных условий моделирования на характеристики нестационарных имитационных процессов / Строганов Д.В., Николаев А.Б., Карташев М.И., Замыцких П.В. // Имитационное моделирование систем управления. – М.: МАДИ, 2012. – С. 90-98.

46. Ким Дж.-О., Мьюллер Ч.У., Клекка У.Р., Олдендерфор М.С., Блэшфилд Р. К. Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989.

47. Козлов О.С., Медведев В.С. Цифровое моделирование следящих приводов. // В кн.: Следящие приводы. В 3-х т. /Под ред. Б.К. Чемоданова. М.: Изд. МГТУ им. Н.Э. Баумана, 1999. Т. 1. С. 711-806.
48. Колесов Ю.Б. Анализ корректности процессов логического управления динамическими объектами // Известия ЛЭТИ. Сб. научн. Трудов / Ленингр. Электротехнич. Ин-т им. В.И.Ульянова (Ленина). – Л.: 1991. – Вып. 436. – с. 65-70.
49. Колесов Ю.Б., Сениченков Ю.Б. Библиотека программ для решения ОДУ. Труды ЛПИ, 462. – СПб.: 1996, с. 116-122.
50. Колесов Ю.Б., Сениченков Ю.Б. Компьютерное моделирование в научных исследованиях и в образовании. "Exponenta Pro. Математика в приложениях", №1, 2003, с. 4-11.
51. Колесов Ю.Б., Сениченков Ю.Б. Программная поддержка активного вычислительного эксперимента В сб. "Научно-технические ведомости СПбГПУ", №1.2004.
52. Колесов Ю.Б., Сениченков Ю.Б.. Визуальное моделирование сложных динамических систем. Изд. «Мир и Семья & Интерлайн», СПб, 2000, 242с.
53. Колесов Ю.Б., Сениченков Ю.Б.. Синхронизация событий при использовании гибридных автоматов для численного моделирования сложных динамических систем. В сб. "Научно-технические ведомости СПбГПУ", №1.2004.
54. Колесов Ю.Б., Цитович И.Г. Имитационная модель участка трикотажного производства // Известия ВУЗ'ов. Технология легкой промышленности, 1993, №6, с.56-61.
55. Колесов Ю.Б., Цитович И.Г. Оценка эффективности новой кругловязальной машины с помощью имитационной модели // Известия ВУЗ'ов. Технология легкой промышленности, 1994, №4, с. 72-77.
56. Красильников В.В. Статистика объектов нечисловой природы. – Наб. Челны: Изд-во Камского политехнического института, 2001. 144 с.

57. Курочкин Е.П., Колесов Ю.Б. Технология программирования сложных систем управления / ВМНУЦ ВТИ ГКВТИ СССР. – М.: 1990. – 112с.
58. Липаев В.В. Надежность программных средств, М.: Синтег, 1998. – 232с.
59. Липаев В.В. Системное проектирование сложных программных средств для информационных систем. - М.: Синтег, 1999. – 224с.
60. Майо Д. С#: Искусство программирования. Энциклопедия программиста: Пер. с англ. СПб.: «ДиаСофтЮП», 2002. – 656 с.
61. Назаров А.И., Сергеев А.В. Система дистанционного контроля знаний в сетях Интернет и Интранет // Дистанционное образование. М. 1999. № 1. С. 11.
62. Никифорова А.М., Попов Д.И. Классификация типов вопросов для адаптивного тестирования // Тезисы докладов V Всероссийской научной конференции студентов и аспирантов «Техническая кибернетика, радиоэлектроника и системы управления». Таганрог. 2000. С. 122 – 123.
63. Никифорова А.М., Попов Д.И. Способы оценки тестовых заданий в системе дистанционного обучения «KnowledgeCT» // Материалы III Всероссийской научной конференции молодых ученых и аспирантов «Новые информационные технологии. Разработка и аспекты применения»: Тезисы докладов. Таганрог, 2000. С. –167 – 168.
64. Никифорова А.М., Попов Д.И., Калашникова Т.Г. Дистанционное образование: тестирование и оценка знаний // VI Междунар. науч.-техн. конф. студентов и аспирантов «Радиоэлектроника, электротехника и энергетика»: Тез. докл. В 3-х т. М., 2000. С. 341–42.
65. Оганесян А.Г. Опыт компьютерного контроля знаний // Дистанционное образование. М. 1999. № 6. С. 30.
66. Огорелков В.И. Основные направления исследования проблем измерения качества знаний учащихся // Проблемы педагогической квалиметрии / Под ред. В.И. Огорелкова. М., 1974. Вып. 1.

67. Орлов А.И. Заводская лаборатория. 1995, Т. 61, № 3.
68. Основы открытого образования / А.А. Андреев, С.Л. Каплан и др.; Отв. ред. В.И. Солдаткин. Т.1. Российский государственный институт открытого образования. М.: НИИЦ РАО, 2002. 676 с.
69. Переверзев В.Ю. Критериально-ориентированное педагогическое тестирование: учебн. пособие. – М.: Логос, 2003.
70. Петров Г.Н. Использование пакета “Model Vision” для создания компьютерных лабораторных работ. // Гибридные системы. Model Vision Studium: Труды междунар. науч.-технич. конф. СПб.: Изд-во СПбГТУ, 2001. – с.53-54.
71. Подчуфаров Ю.Б.. Физико-математическое моделирование систем управления и комплексов / Под ред. А.Г.Шипунова. – М.: Изд-во физико-математической литературы, 2002. – 168с.
72. Попов Д.И. Автоматизация управления процессов аттестации персонала предприятий промышленности: монография. – М.: Изд-во МГУП, 2007. – 178 с.
73. Попов Д.И. Методы и технологии поддержки открытого образования на основе интеллектуальной информационно-образовательной среды дистанционного обучения. / Научное издание. Таганрог: Изд-во ТРТУ, 2003. 168с.
74. Попова Е.Д. Оценка уровня учебных достижений // Известия высших учебных заведений. Проблемы полиграфии и издательского дела №2005, М., 2005.
75. Прицкер А. Введение в имитационное моделирование и язык СЛАМ II: Пер. с англ. - М.: Мир, 1987. - 646с.
76. Проблемы педагогической квалитетрии: Межвуз. сб. тр. / Под ред. В.И. Огорелкова. М., 1973, 1975. Вып. 1, 2; То же / Под ред. В.И. Левина. М., 1984.
77. Самарский А.А., Михайлов А.П.. Математическое моделирование: Идеи. Методы. Примеры. - М.: Наука. Физматлит, 1997.-320 с.

78. Семененко М. Введение в математическое моделирование - М.:Солон-Р, 2002. - 112с.
79. Семенов В.В. Индивидуально-личностный подход в компьютерной технологии тестирования знаний // Аналитические обзоры по основным направлениям развития высшего образования. М. 1998. Вып. 3. С. 49.
80. Состояние и развитие дистанционного образования в мире: Научно-аналитический доклад. М.: Магистр, 1997.
81. Терехов А.Н., Романовский К.Ю, Кознов Дм. В., Долгов П.С., Иванов А.Н. Объектно-ориентированная методология разработки информационных систем и систем реального времени. // Объектно-ориентированное визуальное моделирование / Под ред. Проф. Терехова А.Н. – СПб: Издательство С.-Петербургского университета, 1999. с.4-20.
82. Трудоношин В.А., Пивоварова Н.В.. Математические модели технических объектов - Мн.: Выш. шк.,1988 - 159с.
83. Убиенных Г.Ф., Убиенных А.Г. Сравнительный анализ методов представления знаний в базах знаний. Пенза, Пензенский государственный университет, 2002.
84. Уотермен Д. Руководство по экспертным системам: Пер. с англ. М.: Мир, 1989. 388 с.
85. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие задачи и дифференциально-алгебраические задачи, М., Мир, 1999,- 685с.
86. Чельшкова М.Б. Теория и практика конструирования педагогических тестов: учебное пособие. – М.: Логос, 2002. 432с.
87. Черемных С.В., Семенов И.О., Ручкин В.С. Структурный анализ систем: IDEF-технологии, М.: Финстат, 2001. - 208с.
88. Черных И.В. Simulink: среда создания инженерных приложений. М.: ДИАЛОГ-МИФИ, 2003. – 496с.
89. Шорников Ю.В., Жданов Т.С., Ландовский В.В.. Компьютерное моделирование динамических систем // «Компьютерное моделирование

2003». Труды 4-й межд. научно-техн. конференции, С.Петербург, 24-28 июня 2003г., с.373-380

90. Юдицкий С.А., Покалев С.С.. Логическое управление гибким интегрированным производством // Институт проблем управления. – Препринт. – М., 1989. – 55с.

91. Andersson M. Omola - An Object-Oriented Language for Model Representation, in: 1989 IEEE Control Systems Society Workshop on Computer-Aided Control System Design (CACSD). Tampa, Florida, 1989.

92. Andersson M. OmSim and Omola Tutorial and User's Manual. Version 3.4., Department of Automatic Control, Lund Institute of Technology, 1995, pp.45.

93. Ascher Uri M., Petzold Linda R. Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations. SIAM, Philadelphia, 1998.

94. Avrutin V., Schutz M. Remarks to simulation and investigation of hybrid systems, // Гибридные системы. Model Vision Studium: Труды междунар. науч.-технич. конф. СПб.: Изд-во СПбГТУ , 2001. – с.64-66.

95. Baleani M., Ferrari F., Sangiovanni-Vincentelli A.L., and Turchetti C. HW/SW Codesign of an Engine Management System. In Proc. Design Automation and Test in Europe, DATE'00, Paris, France, March 2000, pp.263-270.

96. Baleani M., Gennari F., Jiang Y., Patel Y., Brayton R.K., and Sangiovanni-Vincentelli A.L.. HW/SW Partitioning and Code Generation of Embedded Control Applications on a Reconfigurable Architecture Platform. In Proc. International Symposium on Hardware/Software Codesign, CODES'02, Estes Park, Colorado, May 2002, pp. 151-156.

97. Booch G., Jacobson I., Rumbaugh J. The Unified Modeling Language for Object-Oriented Development. Documentation Set Version 1.1. September 1997.

98. Borshchev A., Karpov Yu., Kharitonov V. Distributed Simulation of Hybrid Systems with AnyLogic and HLA // Future Generation Computer Systems v.18 (2002), pp.829-839.

99. Brenan K.E., Campbell S.L., Petzold L.R. Numerical solution of initial-value problems in differential-algebraic equations. North-Holland, 1989, 195 p.
100. Bruck D., Elmqvist H., Olsson H., Mattsson S.E. Dymola for multi-engineering modeling and simulation. 2nd International Modelica Conference, March 18-19 2002, Proceedings, pp. 55-1 – 55-8.
101. Bunus P., Fritzson P. Methods for Structural Analysis and Debugging of Modelica Models. 2nd International Modelica Conference, 2002, Proceeding, pp. 157-165.
102. Darnell K., Mulpur A.K.. Visual Simulation with Student VisSim, Brooks Cole Publishing, 1996.
103. Davey, B.A. & Priestley, H.A. Introduction to Lattice and Orders. Cambridge University Press. 1990.
104. Dmitry Popov, Alexander Khadzhinov. "Safety Subsystem of Intelligent Software Complex for Distance Learning" // Proceedings of 2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS 2002), IEEE Inc. 2002. P.464 – 465.
105. Doignon, J-P., Falmagne J-C. (1999) Knowledge Spaces.
106. Esposit J.M., Kumar V., Pappas G.I. Accurate event detection for simulating hybrid systems. Hybrid Systems: Computation and Control, 4th International Workshop, HSCC 2001, Rome, Italy, March 28-30, 2001, Proceedings, pp.204-217.
107. Ferreira J.A., Estima de Oliveira J.P. Modelling hybrid systems using statecharts and Modelica. . In Proc. of the 7th IEEE International Conference on Emerging Technologies and Factory Automation, Barcelona, Spain, 18-21 Oct., 1999, p.1063.
108. Fritzson P., Gunnarson J., Jirstrand M. MathModelica – an extensible modeling and simulation environment with integrated graphics and literate programming/ 2nd International Modelica Conference, March 18-19 2002, Proceedings, pp. 41-54.

- 109.Harel D., Gery E. Executable Object Modeling with Statecharts / Computer, July 1997, pp. 31-42.
- 110.Hyunok Oh, Soonhoi Ha. Hardware-software cosynthesis of multi-mode multi-task embedded systems with real-time constraints. In Proc. International Symposium on Hardware/Software Codesign, CODES'02, Estes Park, Colorado, May 2002, pp. 133-138.
- 111.IMS Content Packaging Information Model, T.Anderson, M.McKell, A.Cooper and W.Young, C.Moffatt, Version 1.1.2. IMS, August 2001.
- 112.IMS Question & Test Interoperability: Overview, C.Smythe, E.Shepherd, L.Brewer and S.Lay, Version 1.2, IMS, September 2001.
- 113.Kesten Y., Pnueli A. Timed and hybrid statecharts and their textual representation. Lec. Notes in Comp. Sci. pp. 591-620, Springer-Verlag, 1992.
- 114.Khartsiev V.E., Shpunt V.K., Levchenko V.F., Kolesov Yu., Senichenkov Yu., Bogotushin Yu. The modeling of synergetic interaction in Theoretical biology. / Tools for mathematical modelling. St. Petersburg, 1999, p.71-73.
- 115.Kolesov Y., Senichenkov Y. A composition of open hybrid automata. Proceedings of IEEE Region 8 International Conference «Computer as a tool», Ljubljana, Slovenia, Sep.22-24,. 2003, v.2, pp. 327-331.
- 116.Koppen, M. Extracting human expertise for constructing knowledge spaces: an algorithm. Journal of Mathematical Psychology, 37, 1993. 1–20.
- 117.Ledin J. Simulation Engineering. CMP Books, Lawrence, Kansas, 2001.
- 118.Mattsson S.E., Elmqvist H., Otter M., Olsson H. Initialization of hybrid differential-algebraic equations in Modelica 2.0. 2nd International Modelica Conference, March 18-19 2002, Proceedings, pp. 9-15.
- 119.Modelica - A Unified Object-Oriented Language for Physical Systems Modeling. Language Specification. Version 2.0, July 10, 2002.
- 120.Modelica - A Unified Object-Oriented Language for Physical Systems Modeling. Tutorial. Version 2.0, July 10, 2002.

121. Modelica – a unified object-oriented language for physical systems modeling. Tutorial. Version 1.4, December 15, 2000.

122. Mosterman P.J. An overview of hybrid simulation phenomena and their support by simulation packages. In Hybrid Systems: Computation and Control '99, vol. 1569 in Lecture Notes in Computer Science, Frits W. Vaandrager and Jan H. van Schuppen (eds.), pp. 165-177, 1999.

123. Mosterman P.J. Hybrid dynamic systems: a hybrid bond graph modeling paradigm and its application in diagnosis. Dissertation for the degree PhD of Electrical Engineering/ - Vanderbilt University, Nashville, Tennessee, 1997.