

На правах рукописи

Жиров Владимир Валерьевич

**РАЗРАБОТКА МЕТОДОВ КВАЛИМЕТРИИ БАЗ ЗНАНИЙ
ПРОМЫШЛЕННЫХ ЭКСПЕРТНЫХ СИСТЕМ
НА ОСНОВЕ ТЕОРИИ НЕАКСИОМАТИЧЕСКОЙ ЛОГИКИ
И ГРАФОАНАЛИТИЧЕСКОГО МОДЕЛИРОВАНИЯ**

Специальность 05.13.01 – Системный анализ, управление
и обработка информации (промышленность)

Автореферат

диссертации на соискание учёной степени
кандидата технических наук

Самара – 2011

Работа выполнена на кафедре информационных технологий Федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Самарский государственный технический университет».

Научный руководитель: доктор технических наук, профессор
Батищев Виталий Иванович

Официальные оппоненты: заслуженный деятель науки РФ,
доктор технических наук, профессор
Дилигенский Николай Владимирович

кандидат технических наук
Петров Максим Владимирович

Ведущая организация: ФГОБУ ВПО «Поволжский государственный университет телекоммуникаций и информатики» (г. Самара)

Защита диссертации состоится 23 декабря 2011 г. в 9-00 на заседании диссертационного совета Д 212.217.03 ФГБОУ ВПО «Самарский государственный технический университет» по адресу: г. Самара, ул. Галактионовская, 141, корпус 6, ауд. 33.

Отзывы по данной работе в двух экземплярах, заверенные печатью, просим направлять по адресу: Россия, 443100, Самара, ул. Молодогвардейская 244, главный корпус, на имя учёного секретаря диссертационного совета Д 212.217.03; факс: (846) 278-44-00.

С диссертацией можно ознакомиться в библиотеке Самарского государственного технического университета (ул. Первомайская, 18)

Автореферат разослан 21 ноября 2011 г.

Учёный секретарь
диссертационного совета Д 212.217.03

Губанов Н.Г.

Общая характеристика работы

Актуальность работы. Управление предприятием в современных условиях связано с контролем постоянно возрастающего числа внутренних и внешних показателей, на изменение которых необходимо своевременно реагировать принятием соответствующих решений. Именно поэтому важным элементом системы управления предприятием являются автоматизированные системы поддержки принятия решений. Высокая сложность и комплексный характер задач управления предприятием требуют использования знаний экспертов из различных предметных областей, что приводит к необходимости применения интеллектуальных систем поддержки принятия решений, построенных на основе технологии экспертных систем (ЭС), и использующих для своего функционирования базы знаний (БЗ).

Высокая скорость изменения среды, в которой функционирует предприятие, и возрастающий объём информации, необходимой для принятия решений, приводят к формированию новых требований к используемым ЭС. Для обеспечения актуальности принимаемых решений в быстро меняющихся условиях необходимо увеличение объёма и повышение частоты обновления БЗ. В то же время, актуальной задачей остаётся оценка соответствия ЭС предъявляемым требованиям. Поскольку соответствие предъявляемым требованиям является основным критерием качества ЭС, указанная задача представляет собой задачу оценки качества, или квалиметрии, ЭС. Новые требования к ЭС порождают новые требования к системам квалиметрии ЭС. Увеличение частоты обновления баз знаний требует более частого проведения квалиметрии, что приводит к необходимости автоматизации процесса квалиметрии. Увеличение объёма баз знаний требует использования методов квалиметрии, эффективных с точки зрения временных затрат на её проведение. Кроме того, направленность новых требований на базы знаний как на ключевой компонент ЭС требует использования методов квалиметрии, учитывающих особенности баз знаний как объекта квалиметрии.

В качестве методологической основы квалиметрии баз знаний ЭС можно рассматривать известные исследования по нескольким направлениям. Первым из таких направлений является общая теория квалиметрии, подробно рассмотренная в работах Г. Г. Азгальдова, А. Г. Варжапетяна, М. М. Калейчика. и др. Общая квалиметрия выделяет основные принципы квалиметрии и предлагает обобщённый алгоритм численной оценки качества. Однако, особенности оценки качества баз знаний и вопросы автоматизации оценки качества остаются за пределами рассмотрения общей квалиметрии. Вторым направлением является оценка качества знаний специалистов, используемая в сфере образования и в сфере управления кадрами. Это направление включает работы А.И. Субетто, А. Н. Майорова, В. С. Аванесова, Д. Н. Шмелёва, Л. М. Спенсер, С. Уидет, и др. В рамках квалиметрии знаний специалистов рассматриваются важные особенности знаний как

объекта квалиметрии, а также методы автоматизации квалиметрии знаний. Однако, специфика этого направления заключается в том, что единственным рассматриваемым носителем знаний является человек. В результате, методы квалиметрии знаний специалистов не могут быть использованы в неизменном виде для оценки качества знаний в технических системах, каковыми являются БЗ ЭС. Третьим направлением является тестирование программного обеспечения, которому посвящены работы Дж. Макгрегора, С. Канера, и др. В рамках этого направления рассматриваются методы квалиметрии программных систем, в том числе методы автоматизации такой квалиметрии. Однако, исследования в сфере тестирования программного обеспечения не учитывают специфику систем, основанных на использовании баз знаний, в результате чего методы тестирования программного обеспечения невозможно использовать в неизменном виде для квалиметрии БЗ ЭС. Таким образом, ни одно из рассмотренных направлений не предлагает готовых к использованию методов оценки качества баз знаний ЭС.

Вместе с тем, исследования в области квалиметрии знаний специалистов показывают, что в сложных предметных областях существует проблема увеличения числа тестов, необходимых для проведения квалиметрии знаний. Увеличение числа тестов, в свою очередь, приводит к увеличению затрат времени и вычислительных ресурсов, задействованных в процессе квалиметрии. Для сокращения таких затрат необходима оптимизация набора проводимых тестов, что требует построения соответствующей модели качества знаний. Перспективными методами, которые можно использовать при построении такой модели, являются методы теории неаксиоматической логики П. Вана, а также методы графоаналитического моделирования.

В связи с вышесказанным представляется актуальной разработка формальных методов автоматизированной квалиметрии баз знаний ЭС, с одной стороны, основанных на положениях общей теории квалиметрии, теории квалиметрии знаний специалистов и теории тестирования программного обеспечения, и с другой стороны, учитывающих специфику оценки качества БЗ ЭС. Также представляется актуальным решение комплекса задач, связанных с сокращением затрат на проведение квалиметрии БЗ ЭС на основе оптимизации набора проводимых тестов.

Целью настоящей диссертационной работы является разработка методов автоматизированной квалиметрии баз знаний экспертных систем, а также методов сокращения затрат на проведение такой квалиметрии путём оптимизации набора проводимых тестов.

Для достижения поставленной цели в работе решаются следующие задачи:

1. Проводится анализ положений общей квалиметрии, обобщённого алгоритма численной оценки качества и общих проблем их применения к оценке БЗ ЭС;
2. Проводится анализ методов квалиметрии, используемых в смежных предметных областях (технической диагностике, контроле качества

программного обеспечения, профессиональном образовании, управлении персоналом) с целью выявления способов решения общих проблем, обнаруженных при оценке БЗ ЭС;

3. Проводится анализ специфики квалиметрии БЗ ЭС и выявляются частные проблемы – неоднозначность при выборе интерфейса взаимодействия с ЭС, а также необходимость сокращения временных и вычислительных затрат на проведение квалиметрии с помощью оптимизации набора проводимых тестов;

4. Проводится морфологический анализ ЭС с целью поиска интерфейсов взаимодействия с системой и сравнительный анализ найденных интерфейсов с точки зрения возможности их применения для квалиметрии БЗ;

5. Формулируются требования к информационному обеспечению процесса квалиметрии БЗ, необходимому для оптимизации набора проводимых тестов;

6. Разрабатывается модель качества знаний, соответствующая информационным потребностям задачи оптимизации состава проводимых тестов;

7. Предлагается методика квалиметрии БЗ ЭС с использованием модели качества знаний, основанная на сравнении результатов обработки системой пользовательских запросов с экспертно заданными эталонными результатами;

8. Предлагаются методы сокращения затрат на проведение квалиметрии БЗ ЭС на основе интервального прогнозирования значений комплексных оценок, а также на основе выявления и использования нетаксономических связей между тестами;

9. На основе предложенных методов сокращения затрат разрабатывается комбинированный алгоритм оптимизации набора проводимых тестов;

10. Рассматривается программная реализация предложенных методов.

Основными методами исследования являются методы системного анализа, графоаналитические методы, методы теории множеств, теории неаксиоматической логики, интервального анализа, «жадные» алгоритмы, метод ветвей и границ, методы объектно-ориентированного программирования.

Научная новизна:

1. Предложена методика квалиметрии баз знаний экспертных систем, основанная на сравнении результатов обработки системой пользовательских запросов с экспертно заданными эталонными результатами, в отличие от существующих, позволяющая проводить автоматизированную оценку качества баз знаний независимо от способа представления знаний внутри системы.

2. Разработана модель качества знаний, отличающаяся от известных моделей использованием направленного графа суждений с представлением истинностных значений суждений с помощью теоретического аппарата неаксиоматической логики П. Вана, позволяющая прогнозировать значения

комплексных оценок качества на основе общего числа тестов и результатов проведённых тестов.

3. Предложены методы уменьшения числа проводимых тестов при квалиметрии БЗ ЭС, отличающиеся от известных интервальным прогнозированием значений комплексных оценок, а также выявлением и использованием нетаксономических связей между тестами, позволяющие сократить временные и вычислительные затраты на проведение квалиметрии знаний.

4. Предложен комбинированный метод оптимизации набора проводимых тестов, отличающийся от известных комплексным использованием нескольких методов уменьшения числа проводимых тестов, а также метода ветвей и границ в сочетании с «жадным» алгоритмом оптимизации, позволяющий сократить временные и вычислительные затраты на проведение квалиметрии знаний.

Практическая значимость заключается в возможности применения полученных методов при решении задач квалиметрии баз знаний промышленных экспертных систем с целью повышения качества разработки таких систем и оценки их применимости для решения новых классов задач.

Реализация и внедрение результатов работы. Разработанное программное обеспечение системы квалиметрии баз знаний было применено в ОАО «ТЯЖМАШ» (г. Сызрань) для мониторинга качества базы знаний, входящей в состав системы поддержки принятия решений по управлению производством. Результаты работы также были использованы в учебном процессе ФГБОУ ВПО «Самарский государственный технический университет» на кафедре информационных технологий. Кроме того, система квалиметрии баз знаний была использована на кафедре педагогики, психологии и психолингвистики ГБОУ ВПО «Самарский государственный медицинский университет» для оценки качества базы знаний экспертной системы «Галатея», предназначенной для контроля профессиональных знаний специалистов в различных предметных областях.

Положения, выносимые на защиту:

1. Методика квалиметрии баз знаний экспертных систем, основанная на сравнении результатов обработки системой пользовательских запросов с экспертно заданными эталонными результатами.
2. Модель качества знаний, основанная на направленном графе суждений с представлением истинностных значений суждений с помощью теоретического аппарата неаксиоматической логики П. Вана.
3. Методы уменьшения числа проводимых тестов при квалиметрии БЗ ЭС, основанные на интервальном прогнозировании значений комплексных оценок, а также на выявлении и использовании нетаксономических связей между тестами.
4. Комбинированный метод оптимизации набора проводимых тестов, отличающийся от известных комплексным использованием нескольких методов уменьшения числа проводимых тестов, а также метода ветвей и границ в сочетании с «жадным» алгоритмом оптимизации.

Апробация работы. Основные положения работы и результаты исследований докладывались и обсуждались на конференциях: Международная научно-техническая конференция «Радиоэлектроника, электротехника и энергетика» (Москва, 2005), Международная научно-техническая конференция «Современные информационные технологии». Весенняя сессия (Пенза, 2005), научные сессии Московского инженерно-физического института (Москва, 2006, 2007, 2008, 2010, 2011), V и VIII Всероссийские научные конференции с международным участием «Математическое моделирование и краевые задачи» (Самара, 2008, 2011), XI Симпозиум «Квалиметрия в образовании: методология, методика, практика» (Москва, 2006), Десятая Всероссийская конференция-семинар «Проектирование, контроль и управление качеством продукции и образовательных услуг» (Москва-Тольятти-Сызрань, 2007), Международная научно-техническая конференция «Синергетика природных, технических и социально-экономических систем» (Тольятти, 2007), Всероссийская научно-методическая конференция «Роль гуманитарных наук в системе современного высшего образования» (Самара, 2008), Всероссийская научно-практическая конференция учёных и педагогов-практиков «Актуальные проблемы развития высшего и среднего образования на современном этапе» (Самара, 2008), Международная научно-практическая конференция «Инновация-2010» (Ташкент, 2010).

Публикации. По теме диссертации опубликовано 20 печатных работ, в том числе 1 статья в журнале из перечня, рекомендованного ВАК РФ, а также получено три свидетельства об официальной регистрации программы для ЭВМ.

Объем и структура диссертационной работы. Диссертация состоит из введения, 4 глав, заключения и выводов, списка использованных источников из 100 наименований, приложений. Работа изложена на 108 страницах, содержит 8 таблиц и 21 рисунок.

Содержание работы

Во введении обоснована актуальность исследуемой проблемы, сформулирована цель и основные задачи работы, аргументирована научная новизна исследований, показана практическая значимость полученных результатов, представлены выносимые на защиту основные положения.

В первой главе проведён анализ положений общей теории квалиметрии, подходов к определению качества и обобщённого алгоритма оценки качества. В результате анализа выявлены основные проблемы применения известных подходов к оценке качества знаний. Первая проблема заключается в том, что знания не являются материальным объектом, и прямое измерение их свойств невозможно. В результате возникает задача поиска объекта оценки, который позволял бы проводить прямые измерения свойств, и свойства которого были бы связаны со свойствами знаний известной зависимостью. Вторая проблема заключается в неочевидности набора свойств знаний. Для её решения требу-

ется выявление свойств знаний, определяющих качество, которые могут быть косвенно измерены или оценены. Третья проблема заключается в отсутствии объективного, очевидного эталона знаний. В результате требуется поиск таких способов определения базовых значений выявленных свойств, и определение соответствующей функции оценки качества.

Для решения возникших проблем был проведён анализ методов квалиметрии знаний в сферах образования, управления персоналом, методов оценки качества программного обеспечения в сфере информационных технологий, а также анализ частных методов квалиметрии в сфере технической диагностики. В результате такого анализа были выявлены общие закономерности, свойственные методам квалиметрии во всех тех сферах, где объектом оценки являются знания:

- оценка знаний в этих сферах происходит путём оценки деятельности, поскольку именно в процессе деятельности происходит проявление знаний;
- деятельность рассматривается как преобразование некоторых входных данных в выходные данные (тестовое задание – ответ, ситуация – поведение, контрольное логическое утверждение – результат проверки);
- деятельность имеет сложную структуру и подвергается функциональной декомпозиции на множество элементарных действий, каждое из которых также характеризуется входными и выходными данными;
- декомпозиция деятельности проводится таким образом, чтобы каждое действие могло быть оценено с точки зрения соответствия выходных данных целям действия, и чтобы такая оценка была двузначной («соответствует» или «не соответствует», «истина» или «ложь»);
- процедуры оценки соответствия, состав и вид используемых эталонов различны для разных действий, однако декомпозиция деятельности организована таким образом, чтобы число типов используемых процедур было конечным, и, по возможности, минимальным;
- помимо функциональной декомпозиции деятельности, проводится смысловая декомпозиция знаний предметной области с образованием таксономической структуры знаний;
- результаты оценки действий сопоставляются с таксонами, что позволяет использовать связи между таксонами для агрегации оценок и расчёта комплексной оценки.

В соответствии с выявленными закономерностями структурная схема компонентов, задействованных в процессе оценки знаний может быть изображена следующим образом (рисунок 1).

После выделения компонентов, задействованных в процессе квалиметрии знаний необходимо было выявить эквивалентные им компоненты, входящие в состав систем поддержки принятия решений. Для этого был проведён морфологический анализ существующих интеллектуальных ЭС. В связи с труднодоступностью конкретных реализаций ЭС, в качестве исходных данных для анализа использовались шесть обобщённых структурных схем интеллектуальных систем, предложенных в разное время

(1979 – 2010) разными авторами. В результате анализа были составлены два собственных варианта обобщённой структурной схемы ЭС, выделены интерфейсы взаимодействия ЭС со средой, произведена классификация этих интерфейсов и сделан вывод о целесообразности оценки ЭС через интерфейс диалога с пользователем.

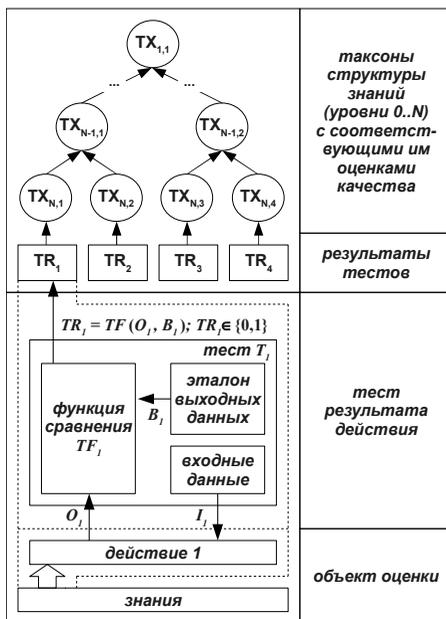


Рисунок 1 – Схема компонентов, задействованных в процессе оценки знаний

В этом случае действия, подвергаемые непосредственной оценке, заключаются в обработке ЭС запроса пользователя, входными данными I являются собственно запросы, выходными данными O – ответы ЭС на запросы, при этом аргументами функции оценки TF являются ответы ЭС и эталонные значения B ответов или их частей.

Детализация оценок действий и таксономической структуры знаний требовала обращения к основным целям квалиметрии знаний в БЗ ЭС:

1. оценка возможности применения ЭС при заданных требованиях к принимаемым решениям;
2. выявление недоработок в ЭС и планирование мероприятий по их устранению.

Поскольку возможность применения ЭС определяется количеством и характером недоработок, для достижения поставленных целей необходимо выполнение следующих задач:

- классификации недоработок ЭС;
- определение связи между классами недоработок и оценками действий;
- определение допустимого числа недоработок для каждого класса недоработок.

При этом выбор критериев классификации недоработок обусловлен информативностью этих критериев для локализации и устранения недоработок. Например, в качестве критериев могут использоваться:

- связь недоработки с определённым фрагментом БЗ ЭС;
- связь недоработки с определённым этапом формирования БЗ ЭС;
- связь недоработки с определённым видом принимаемых решений;
- степень негативного влияния недоработки на процесс принятия решений.

Таким образом, при оценке знаний в БЗ ЭС, целесообразен переход от использования структуры знаний к использованию структуры недоработок БЗ в качестве основы для системы комплексных оценок. Поскольку структура базы знаний не зависит от результатов проведённых тестов, изменение формулировки со «структуры показателей качества» на «структуру недоработок» в данном случае не имеет значения.

С учётом приведённых выше примеров критериев классификации недоработок, можно выделить следующие особенности структуры недоработок:

- наличие нескольких критериев классификации;
- возможность многоуровневой классификации по отдельным критериям с образованием таксономической структуры классов;

Такие структурные особенности приводят к значительному росту числа классов при увеличении сложности оцениваемой ЭС. Поскольку оценки классов недоработок являются комплексными и основаны на дифференциальных оценках действий, увеличение числа классов недоработок приводит к увеличению необходимого числа дифференциальных оценок. В результате число действий, подвергаемых непосредственной оценке оказывается достаточно большим, чтобы поставить проблему сокращения временных и вычислительных затрат на проведение оценки.

Во второй главе рассмотрена модель качества знаний, предназначенная для использования в процессе квалиметрии знаний в БЗ ЭС. Требования к модели были разделены на две группы:

1. требования, обусловленные особенностями квалиметрии знаний;
2. требования, обусловленные информационными потребностями алгоритма оптимизации набора проводимых тестов.

Требования, обусловленные особенностями квалиметрии знаний включают:

- использование в качестве дифференциальных оценок множества результатов тестов $TRS = \{ TR_i | i = \overline{1..N} \}$; $TR_i \in \{0, 1\}$, полученных сравнением выходных данных системы $\{ O_i | i = \overline{1..N} \}$ с эталонными выходными данными $\{ B_i | i = \overline{1..N} \}$ при заданных входных данных $\{ I_i | i = \overline{1..N} \}$.

- представление результата в виде множества QRS комплексных оценок, соответствующих классам недоработок БЗ ЭС;
- использование в качестве основы для расчёта комплексных оценок многокритериальной классификации недоработок БЗ ЭС, сформированной экспертами;

Требования, обусловленные информационными потребностями алгоритма оптимизации набора проводимых тестов включают:

- возможность расчёта текущих значений комплексных оценок на основе любого подмножества результатов тестов;
- возможность интервального прогнозирования окончательных значений комплексных оценок на основе текущих значений;
- возможность пересчёта текущих значений комплексных оценок и уточнения прогнозов по мере пополнения множества результатов тестов.

Основой модели является направленный ациклический граф оценок качества $G = \langle GV, GE \rangle$, включающий три типа вершин GV (рисунок 2):

1. истоки S , соответствующие дифференциальным оценкам действий Q_S ;
2. стоки R , соответствующие комплексным оценкам по классам недоработок QRS ;
3. промежуточные вершины A , соответствующие комплексным оценкам по подклассам недоработок.

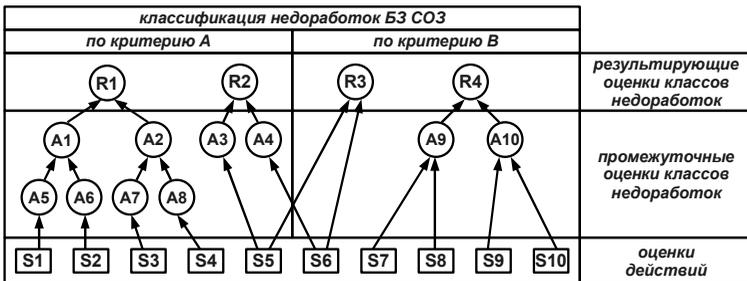


Рисунок 2 – Направленный граф оценок качества

В зависимости от необходимой степени детализации оценки, вершины из множества A могут переходить в множество R , и наоборот, при соблюдении условия $R \neq \emptyset$. Состав множества S постоянен и определяется набором имеющихся тестов для оценки действий. Вне зависимости от типа, каждой вершине V графа соответствует суждение J_V о максимальном значении оценки. Под суждением в данном случае подразумевается высказывание, для которого можно оценить степень его истинности. Для представления истинностных значений суждений используется подход, предложенный П. Ваном в работах, посвящённых теоретическому аппарату неаксиоматической логики. В качестве меры истинности суждения используется число свидетельств (weight of evidence) в поддержку или в опровержение этого суждения,

обозначаемые соответственно w_+ и w_- . Общее число свидетельств для суждений определено как $w = w_+ + w_-$, и для представления информации о свидетельствах достаточно любых двух из трёх рассмотренных величин. Значение и достоверность оценки выражены функциями от числа свидетельств, предложенными П. Ваном – частотой суждения f и уверенностью в суждении c соответственно:

$$Q_V = f(J_V) = \frac{w_+(J_V)}{w(J_V)} = \frac{w_{V+}}{w_V} \quad (1)$$

$$C_V = c(J_V) = \frac{w(J_V)}{w(J_V) + k} = \frac{w_V}{w_V + k}, \quad k = const, \quad k \in [1, \infty] \quad (2)$$

Начальное число всех типов свидетельств для всех вершин равно нулю. С каждой вершиной $V \in S$ связан тест T_V , положительный результат которого увеличивает на единицу число свидетельств w_{V+} , а отрицательный результат – число свидетельств w_{V-} .

Дуги GE графа оценок качества однотипны и используются для перераспределения свидетельств между вершинами. При перераспределении свидетельств число свидетельств, соответствующее вершине V рассчитывается как сумма свидетельств, соответствующих вершинам, непосредственно связанным с V входящими в V дугами $E \in GE$:

$$w_V = \sum_{X: E = \{X, V\}} w_X \quad (3)$$

Указанный способ расчёта используется для всех типов свидетельств – как в поддержку, так и в опровержение J_V . Использование понятия числа простых путей $NPT(X, V)$ от вершины X до вершины V позволяет записать формулу перераспределения свидетельств между истоками и стоками графа G в следующем виде:

$$w_V = \sum_{X \in S} w_X \cdot NPT(X, V) \quad (4)$$

По сравнению с неаксиоматической логикой П. Вана, где отсутствуют ограничения на число свидетельств относительно суждений, и $w \in [0; \infty)$, при квалиметрии баз знаний с помощью тестов максимальное число свидетельств w_{max} ограничено количеством доступных тестов:

$$w_{max} = \begin{cases} 1, & V \in S \\ \sum_{X: E = \{X, V\}} w_{Xmax}, & V \in A \cup R \end{cases} \quad (5)$$

Такое ограничение позволяет на любом этапе ST_i процесса квалиметрии на основе текущих значений числа свидетельств w_{V+} и w_V сделать интервальный прогноз $[wl_{V+}; wh_{V+}]$ окончательного значения числа свидетельств, которое будет получено по завершении процесса квалиметрии на этапе STM (рисунок 3, а и б).

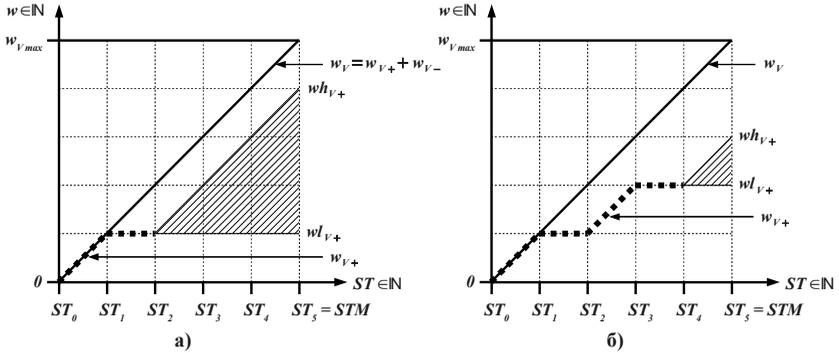


Рисунок 3 – Графическая интерпретация прогнозируемых значений числа свидетельств

Применение формулы (1) к значениям границ интервала $[wl_{V+}; wh_{V+}]$ позволяет получить интервальный прогноз $[QL_V; QH_V]$ значения оценки Q_V :

$$QL_V = \frac{w_{V+}}{w_{Vmax}}; \quad QH_V = \frac{w_{V+} + w_{Vmax} - w_V}{w_{Vmax}} \quad (6)$$

В третьей главе рассмотрены методы сокращения затрат на проведение квалиметрии баз знаний с помощью тестов. В первую очередь, были выделены две группы таких методов сокращения затрат:

1. методы, основанные на уменьшения затрат на проведение тестов;
2. методы, основанные на уменьшения числа фактически проводимых тестов.

Использование первой группы методов затруднено, поскольку затраты на выполнение запросов внутри ЭС не могут быть компенсированы системой оценки, а повышение эффективности входящих в состав теста процедур оценки определяется выбранным интерфейсом взаимодействия с конкретной ЭС, и потому не имеет общей реализации, применимой к любой ЭС.

В то же время, использование методов второй группы, основанных на уменьшении числа фактически проводимых тестов позволяет получить такую общую реализацию. В этой группе были выделены два метода:

1. метод, основанный на прогнозировании значений комплексных оценок;
2. метод, основанный на выявлении и использования нетаксономических связей между тестами.

Метод, основанный на прогнозировании значений комплексных оценок предполагает, что целью квалиметрии знаний является отнесение объекта оценки – БЗ ЭС – к одному из заранее определённых классов качества на основе значений комплексных оценок. При этом для определения фактического класса качества QCA БЗ ЭС используется цепочка расчётов «значения оценок» – «частные значения классов качества по значениям оценок» – «фактический класс качества»: $\{Q_V | V \in R\} \rightarrow \{QCAF_V | V \in R\} \rightarrow QCA$. С помощью рассмотренной в главе 3 модели качества знаний, на основе результатов проведённых тестов возможно интервальное прогнозирование значения комплексных оценок, заключающееся в определении интервала QF_V , гарантированно включающего точное значение оценки Q_V . Это позволяет заменить цепочку расчётов фактического класса качества на новую: «прогнозируемые интервалы значений оценок» – «прогнозируемые интервалы частных значений класса качества» – «прогнозируемый интервал фактического класса качества»: $\{QF_V | V \in R\} \rightarrow \{QCAF_V | V \in R\} \rightarrow QCAF$, $QCAF = [QCAL, QCAN]$. Такая замена осуществляется применением формул первоначальной цепочки расчётов к границам прогнозируемых интервалов. Особенности расчёта фактического класса качества обеспечивают возможность схождения интервала $QCAF$ в точку прежде, чем сойдутся в точку все интервалы значений оценок $\{QF_V | V \in R\}$. Таким образом, выполнение условия $QCAL = QCAN$ в процессе квалиметрии позволяет сделать вывод о фактическом классе качества БЗ ЭС и не проводить тесты, результаты которых ещё неизвестны.

Метод, основанный на выявлении и использования нетаксономических связей между тестами предполагает, что в конкретных случаях квалиметрии между результатами отдельных тестов существует взаимосвязь, которая может быть выявлена экспертами при формировании тестов и позволяет на основании положительного или отрицательного результата одного теста сделать вывод о результате другого, связанного с первым, теста. В случае наличия такой взаимосвязи, целесообразна проверка связей между тестами в процессе квалиметрии знаний, и исключение из множества запланированных к проведению тестов тех, результаты которых можно прогнозировать на основе результатов ранее проведённых тестов.

Сочетание рассмотренных методов сокращения затрат на проведение квалиметрии знаний позволило разработать комбинированный метод оптимизации набора проводимых тестов. При этом использовалась следующая постановка задачи оптимизации. Даны:

1. множество тестов $TS = \{T_V | V \in S\}$;
2. множество классов качества БЗ ЭС $QCS = \{QC_i | i = \overline{1, QCN}\}$;
3. граф G , определяющий связь между множеством результатов тестов TRS и множеством прогнозируемых интервалов значений комплексных оценок $QFS = \{QF_X | X \in R\}$;

4. множество классифицирующих интервалов $QCRS$ классов качества для всех комплексных оценок – $QCRS = \{QCR_X^i | i=1, \overline{QCN}, X \in R\}$ – определяющее связь между прогнозируемыми значениями QFS комплексных оценок и прогнозируемыми значениями $QCAL$ и $QCAH$ фактического класса качества QCA .

Каждый тест T_V имеет результат TR_V , связанный с прогнозируемыми значениями QFS комплексных оценок формулами перераспределения свидетельств в соответствии с конфигурацией графа G , а также формулами расчёта прогнозируемых значений оценки. На основе прогнозируемых значений QFS комплексных оценок и границ классифицирующих интервалов $QCRS$ определяются прогнозируемые значения $QCAL$ и $QCAH$ фактического класса качества QCA БЗ ЭС. Каждому тесту T_V соответствует известная величина абсолютных затрат $TAC(T_V)$. Необходимо найти такое упорядоченное подмножество тестов $ATS = \{AT_j | j=1, \overline{ATN}\}$, $AT_j \in TS$, проведение которого обеспечит равенство $QCAL = QCAH$ при наименьших суммарных абсолютных затратах $ATSAC$ на проведение тестов, где

$$ATSAC = \sum_{j=1}^{ATN} TAC(AT_j) \quad (7)$$

Решение поставленной задачи оптимизации затрудняется тем, что результат TR_V любого теста T_V становится известен только после проведения этого теста. Вследствие этого, прогнозируемые значения комплексных оценок и прогнозируемые значения фактического класса качества уточняются после проведения каждого теста. Таким образом, информация, необходимая для поиска глобального оптимума становится доступной в полном объёме только по окончании процесса квалиметрии, когда множество ATS тестов уже известно и не подлежит изменению.

Преодоление указанной трудности возможно путём использования для решения задачи оптимизации «жадного» алгоритма. Для этого решение разбивается на последовательность шагов, и на каждом шаге делается локально оптимальный выбор, при этом предполагается, что итоговое решение также будет оптимальным. На каждом шаге $j=1, \overline{ATN}$ алгоритма производится выбор очередного проводимого теста AT_j из множества всех возможных вариантов APT . Множество APT представляет собой множество TS , из которого исключены тесты с известными результатами:

$$APT = TS \setminus \{T_V | V \in S, |TRS_V| + |TRAS_V| \geq NTR_V\} \quad (8)$$

Результаты тестов, исключаемых из множества APT могут быть известны как за счёт проведения этих тестов, так и за счёт косвенного определения на основе частичных связей между результатами тестов.

После определения множества возможных вариантов очередного теста APT для поиска оптимального варианта используется метод ветвей и границ. Ветвление заключается в разделении множества APT на три подмножества тестов: неактуальных BPT , актуальных MPT , и приоритетных GPT :

$$\begin{aligned}
 BPT &= \{T_V \mid T_V \in APT, \forall Q_X \in QS(T_V) \ Q_X \in QBS\} \\
 GPT &= \{T_V \mid T_V \in APT, \exists Q_X \in QS(T_V): Q_X \in QGS\} \\
 MPT &= APT \setminus (BPT \cup GPT) \\
 QS(T_V) &= \{Q_X \mid X \in R, NPT(V, X) > 0\} \\
 QBS &= \{Q_X \mid X \in R, QCAL_X = QCAH_X\} \\
 QGS &= \{Q_X \mid X \in R, QCAL_X \neq QCAH_X, QCAL \in [QCAL_X; QCAH_X]\}
 \end{aligned} \tag{9}$$

Тесты, принадлежащие множеству BPT гарантированно не входят в оптимальное решение, поскольку их проведение не изменит значений $QCAL$ и $QCAH$. Тесты принадлежащие множеству MPT способны изменить только верхнюю границу $QCAH$ прогнозируемого интервала значений фактического класса качества. При этом, проведение тестов только из множества MPT не позволяет обеспечить выполнение целевого равенства $QCAL = QCAH$. В то же время, проведение тестов из множества GPT , в зависимости от результатов тестов, способно изменить обе границы $QCAL$ и $QCAH$ прогнозируемого интервала значений фактического класса качества, и также способно обеспечить выполнение целевого равенства. Таким образом, по крайней мере один тест из множества GPT гарантированно входит в оптимальное решение, поскольку без проведения тестов из множества GPT невозможно получение любого решения, в том числе оптимального. Множество MPT не обладает таким свойством, оно может не содержать ни одного теста, входящего в оптимальное решение. На основе вышесказанного можно считать множество GPT предпочтительным при выборе очередного теста.

В качестве критерия при поиске локально оптимального решения в множестве GPT целесообразно использовать величину удельных затрат на проведение теста. Величина удельных затрат $TSC(T_V)$ определяет размер абсолютных затрат, приходящийся на единицу квалиметрической ценности $TW(T_V)$ теста T_V :

$$TSC(T_V) = \frac{TAC(T_V)}{TW(T_V)} \tag{10}$$

Квалиметрическая ценность теста $TW(T_V)$ определяет «степень принадлежности» теста T_V множеству GPT , и основывается на количестве простых путей от вершины V до всех вершин, связанных с оценками, изменение которых способно изменить обе границы $QCAL$ и $QCAH$, и обеспечить выполнение целевого равенства $QCAL = QCAH$:

$$TW(T_V) = \sum_{X: Q_x \in QGS} NPT(V, X) \quad (11)$$

С использованием приведённых определений квалиметрической ценности и величины удельных затрат, целесообразно выбирать в качестве очередного такой тест, принадлежащий множеству GPT , удельные затраты на проведение которого минимальны.

В четвёртой главе рассмотрена программная реализация системы автоматизированной квалиметрии баз знаний, приведены алгоритмы предложенных методов оптимизации набора проводимых тестов, а также структуры данных, используемые для представления модели качества знаний в рамках парадигмы объектно-ориентированного программирования. Разработана архитектура системы квалиметрии баз знаний, структурная схема которой представлена на рисунке 4.



Рисунок 4 – Структурная схема системы квалиметрии знаний

Анализ разработанных методов оптимизации состава проводимых тестов, а также эксперименты, проведённые с использованием программной реализации показали, что эффективность методов в значительной степени определяется конфигурацией классифицирующих интервалов классов качества и фактическими показателями качества системы. В конкретных случаях

квалиметрии было достигнуто снижение затрат на проведение тестов на 24 процента от первоначального уровня затрат, без снижения достоверности оценок качества.

В заключении приведены основные научные результаты работы.

Основные результаты и выводы

1. На основе анализа положений общей квалиметрии и обобщённого алгоритма численной оценки качества выявлены общие проблемы их применения к оценке знаний в БЗ ЭС.
2. На основе анализа методов квалиметрии, используемых в смежных предметных областях (профессиональном образовании, управлении персоналом, технической диагностике и контроле качества программного обеспечения) выявлены способы решения общих проблем, связанных с применением общей квалиметрии и обобщённого алгоритма численной оценки качества к оценке знаний в БЗ ЭС;
3. На основе анализа специфики квалиметрии знаний в БЗ ЭС выявлены частные проблемы, возникающие в процессе квалиметрии – неоднозначность при выборе интерфейса взаимодействия с ЭС, а также необходимость сокращения временных и вычислительных затрат на проведение оценки знаний с помощью оптимизации набора проводимых тестов.
4. На основе морфологического анализа ЭС и сравнительного анализа найденных интерфейсов взаимодействия сформулированы критерии выбора интерфейса для организации взаимодействия с системой в процессе оценки знаний.
5. Сформулированы требования к информационному обеспечению процесса квалиметрии БЗ, необходимому для оптимизации набора проводимых тестов;
6. Разработана модель качества знаний, соответствующая информационным потребностям задачи оптимизации набора проводимых тестов;
7. Предложена методика квалиметрии БЗ ЭС с использованием модели качества знаний, основанная на сравнении результатов обработки системой пользовательских запросов с экспертно заданными эталонными результатами;
8. Предложены методы сокращения затрат на проведение квалиметрии БЗ ЭС на основе прогнозирования значений комплексных оценок, а также на основе выявления и использования нетаксономических связей между тестами;
9. На основе предложенных методов сокращения затрат разработан комбинированный метод оптимизации набора проводимых тестов;
10. Рассмотрена программная реализации предложенных методов.

Опубликованные работы по теме диссертации

Публикации в изданиях из перечня ВАК:

1. Автоматизация решения задачи расстановки кадров предприятия на основе системы компетенций [Текст] / В.В. Жиров // Вестн. Сам. гос. техн. ун-та. Сер.: Технические науки. – Самара, 2010. – №2 (26). – С.74-79

Публикации в других изданиях:

2. Разработка модели качества знаний с учётом динамики процесса квалиметрии [Текст] / В.И. Батишев, В.В. Жиров // Математическое моделирование и краевые задачи. Труды восьмой Всероссийской научной конференции с международным участием. – Самара: СамГТУ, 2011. – С.130-134
3. Автоматизированная квалиметрия баз знаний на основе тестирования программного обеспечения систем поддержки принятия решений [Текст] / В.В. Жиров // Научная сессия НИЯУ МИФИ-2011. Аннотации докладов. – М.: МИФИ, 2011. – Т.3. – С.69
4. Система оценки знаний с использованием нечетких представлений [Текст] / В.Г. Жиров, В.В. Жиров // Научная сессия МИФИ-2006. Сб. научных трудов. – М.: МИФИ, 2006. – Т.3. – С.148-149
5. Проектирование системы оценки знаний с нечетким представлением информации [Текст] / В.Г. Жиров, В.В. Жиров // Научная сессия МИФИ-2007. Сб. научных трудов. – М.: МИФИ, 2007. – Т.3. – С.90-91
6. Оптимизация объёма модели знаний предметной области на ограниченном множестве семантик [Текст] / В.В. Жиров // Математическое моделирование и краевые задачи. Труды пятой Всероссийской научной конференции с международным участием. Ч. 4: Информационные технологии в математическом моделировании. – Самара: СамГТУ, 2008. – С.51-54
7. Способы организации адаптивного тестирования в автоматизированной системе оценки качества знаний [Текст] / А.Н. Краснов, В.В. Жиров // Международная научно-практическая конференция «Инновация-2010». Сборник научных статей. – Ташкент, 2010. – С.283-284
8. Распределенный учебный программный комплекс [Текст] / В.В. Жиров, Т.М. Егорова // Радиоэлектроника, электротехника и энергетика. Двенадцатая международная научно-техническая конференция студентов и аспирантов. Тезисы докладов. Том 1. – М., 2005. – С.438-439
9. Комплекс программных средств для создания электронных учебных пособий [Текст] / Т.М. Егорова, В.В. Жиров // Современные информационные технологии. Сборник статей международной научно-технической конференции. Выпуск 2. – Пенза, 2005. – С.138-139

10. Объединение баз знаний [Текст] / В.Г. Жиров, В.В. Жиров // Научная сессия МИФИ-2008. Сб. научных трудов . – М.: МИФИ, 2008. – Т.10: Интеллектуальные системы и технологии . – С.107-108
11. Применение базы знаний для диагностики состояния производственного комплекса [Текст] / В.Г. Жиров, В.В. Жиров // Научная сессия НИЯУ МИФИ-2009. Аннотации докладов . – М.: МИФИ, 2009. – Т.3: Нанотехнологии и нанотехнологии . – С.66
12. Применение базы знаний для оценки состояния однородных объектов [Текст] / В.Г. Жиров, В.В. Жиров // Научная сессия НИЯУ МИФИ-2010. Аннотации докладов . – М.: МИФИ, 2010. – Т.3 . – С.83
13. Формирование стандартной языковой картины предметной области как основа теоретической подготовки студентов [Текст] / А.Н. Краснов, О.Н. Моисеева, Е.А. Слоева, А.Н. Меркушев, С.А. Никаев, В.В. Жиров, И.С. Мошков // Актуальные проблемы развития высшего и среднего образования на современном этапе: Материалы Всероссийской научно-практической конференции ученых и педагогов-практиков. Т.1. – Самара: Изд-во Самарского научного центра РАН, 2008. – С.271-273
14. Некоторые проблемы разработки традиционной тестовой технологии для диагностики качества теоретической подготовки студентов медицинского вуза по клиническим дисциплинам [Текст] / А.Н. Краснов, В.В. Жиров, Я.А. Краснов, А.А. Юрин // XI Симпозиум «Квалиметрия в образовании: методология, методика, практика». Сборник №2 «Общие проблемы квалиметрии в образовании» Часть 1. – М.: Исследовательский центр проблем качества подготовки специалистов, 2006. – С.18-29
15. ЭС «ГАЛАТЕЯ» как прототип технологии контроля теоретических знаний ближайшего будущего [Текст] / А.Н. Краснов, А.Н. Меркушев, С.А. Никаев, В.В. Жиров, Т.С. Краснова // Проектирование, контроль и управление качеством продукции и образовательных услуг. Материалы десятой Всероссийской конференции-семинара. Тезисы докладов. Часть I. – Москва–Тольятти–Сызрань, 2007. – С.236-240
16. Отдаленная эффективность когнитивного подхода к обучению при его применении в медицинском ВУЗе [Текст] / А.Н. Краснов, А.Ф. Амиров, О.Н. Моисеева, Е.Е. Липатова, Т.С. Краснова, В.В. Жиров // Проектирование, контроль и управление качеством продукции и образовательных услуг. Материалы десятой Всероссийской конференции-семинара. Тезисы докладов. Часть I. – Москва–Тольятти–Сызрань, 2007. – С.224-228
17. Понятийный подход к управлению качеством образовательного процесса в медицинском университете [Текст] / А.Н. Краснов, О.Н. Моисеева, А.Н. Меркушев, Т.С. Краснова, В.В. Жиров // Проектирование, контроль и управление качеством продукции и образовательных услуг. Материалы десятой Всероссийской конференции-семинара. Тезисы докладов. Часть I. – Москва–Тольятти–Сызрань, 2007. – С.228-232

18. Ближайшая эффективность синергетического подхода при его применении в образовательном процессе в медицинском университете [Текст] / А.Н. Краснов, А.Ф. Амиров, О.Н. Моисеева, Е.Е. Липатова, Т.С. Краснова, В.В. Жиров // Синергетика природных, технических и социально-экономических систем. Сборник статей Международной научно-технической конференции. Часть 1. – Тольятти, 2007. – С.74-77
19. Проблемы разработки тестовых заданий открытого типа для оценки профессиональных знаний [Текст] / В.В. Жиров // Роль гуманитарных наук в системе современного высшего образования: материалы Всероссийской научно-методической конференции. – Самара: ООО «МатриКС», 2008. – С.87-89
20. Проблемы совместного использования тестовых заданий открытого и закрытого типов при оценке профессиональных знаний [Текст] / В.В. Жиров // Роль гуманитарных наук в системе современного высшего образования: материалы Всероссийской научно-методической конференции. – Самара: ООО «МатриКС», 2008. – С.89-91

Программы, зарегистрированные в Реестре программ для ЭВМ:

21. Программный комплекс для формирования электронных и мультимедийных учебников и выборки данных, соответствующих тематическому запросу пользователя («Ментор-1»). – Свидетельство об официальной регистрации программы для ЭВМ № 2005611571, заявка № 2005611008 от 11 мая 2005 г. Зарегистрировано в Реестре программ для ЭВМ 27 июня 2005 г. / А.Н. Краснов, Н.Ю. Куваев, А.В. Дидик, В.В. Жиров, А.П. Журавлев, А.А. Юрин
22. Программный комплекс для контроля уровня знаний в электронных и мультимедийных учебных пособиях гуманитарного и естественно-научного типов («Квестор-1»). – Свидетельство об официальной регистрации программы для ЭВМ № 2006612899, заявка № 2006611939 от 13 июня 2006 г. Зарегистрировано в Реестре программ для ЭВМ 11 августа 2006 г. / А.Н. Краснов, Н.Ю. Куваев, В.В. Жиров, А.А. Юрин
23. Программа для оценки знаний в различных предметных областях по ответам на набор вопросов («Квестор-2»). – Свидетельство об официальной регистрации программы для ЭВМ № 2008615267, заявка № 2008614286 от 15 сентября 2008 г. Зарегистрировано в Реестре программ для ЭВМ 31 октября 2008 г. / И.Л. Давыдкин, А.Н. Краснов, Н.Ю. Куваев, В.В. Жиров

Автореферат отпечатан с разрешения диссертационного совета Д212.217.03
ФГБОУ ВПО «Самарский государственный технический университет»
(протокол № 6 от 16 ноября 2011 г.)

Заказ №1115. Тираж 100 экз.

Отпечатано в типографии.
ФГБОУ ВПО «Самарский государственный технический университет»
Отдел типографии и оперативной печати
443100 г. Самара ул. Молодогвардейская, 244, корпус 8