

УДК 519.712.2

В.В. Курейчик, П.В. Сороколетов, С.Н. Щеглов**АНАЛИЗ СОВРЕМЕННОГО СОСТОЯНИЯ АВТОМАТИЗИРОВАННЫХ СИСТЕМ ПРИОБРЕТЕНИЯ И ПРЕДСТАВЛЕНИЯ ЗНАНИЙ***

В настоящее время в механизмах автоматизированного приобретения знаний можно выделить две основные группы. Это автоматизированные системы приобретения знаний от экспертов с использованием множества наработанных технологий и методологий [1]. И новое направление – автоматизированные системы извлечения знаний без участия экспертов из существующих знаний и из текстов, массивов данных (Data mining, Knowledge Discovery in Databases) [2, 3]. Представители обоих направлений широко известны программными средствами, первое поколение которых появилось еще во второй половине 80-х годов. Это SIMER+MIR, OPAL, TEIRESAIS, АРИАДНА [4, 5]. Второе поколение, начиная с 90-х годов [6] представлено широким применением методологии KADS, в основе которой лежит понятие интерпретационной модели, позволяющей рассматривать процессы структурирования и формализации знаний как «интерпретацию» лингвистических знаний в другие представления и структуры. Представитель, данной методологии оболочка CommonKADS, KELLY, MEDIS [6, 7]. Третье поколение – это более сложные, гибкие, использующие графические возможности современных рабочих станций, достижения программных средств и CASE-технологий [8], представлено в [1].

Существуют различные классификации систем приобретения знаний – по выразительности и мощности инструментальных средств, по обобщенным характеристикам, на основе функционального подхода, интегрированная классификация [5-9]. Исходя из новых тенденций в инженерии знаний, можно предложить следующую схему систем приобретения знаний, представленную на рис. 1.

Обобщая современное состояние программных средств приобретения знаний, можно выявить две группы проблем, характерных для этих систем [6]:

а) методологические проблемы:

- ◆ слабая проработанность;
- ◆ отсутствие единого теоретического базиса теоретических аспектов процессов извлечения знаний; процедуры структурирования знаний;
- ◆ жесткость моделей представления знаний;
- ◆ несовершенство математического базиса моделей представления знаний;
- ◆ эмпиричность процедуры выбора программного инструментария и процесса тестирования;

б) технологические проблемы:

- ◆ отсутствие концептуальной целостности и согласованности между отдельными приемами и методами инженерии знаний;
- ◆ недостаток квалифицированных специалистов в области инженерии знаний;
- ◆ неполнота методов структурирования знаний, отсутствие классификации и рекомендаций по выбору подходящего метода;
- ◆ жесткость программных средств, их низкая адаптивность.

Несмотря на имеющиеся недоработки в методологии и технологии, автоматизированные системы знаний позволяют расширять рынок ЭС [9]. Улучшение в

* Работа выполнена при поддержке РФФИ (гранты № 07-01-00174, № 08-01-00473).

системах эргономических показателей, использование знаний о процессе извлечения и преобразования знаний расширяют возможности автоматизации процесса накопления знаний. Перспективным направлением является разработка систем, которые имеют банки данных, машины поиска знаний из данных.

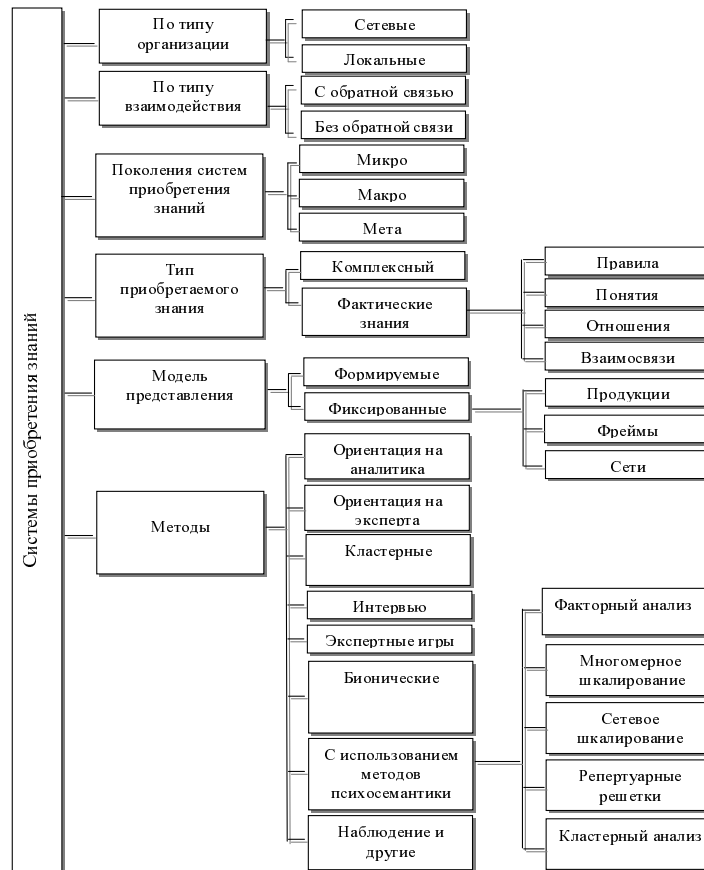


Рис. 1. Классификация систем приобретения знаний

Анализ моделей представления знаний. Существует множество моделей представления знаний. Практически их можно свести к четырем типичным моделям представления знаний [6, 10, 11]:

- ◆ продукционная модель;
- ◆ логическая модель;
- ◆ фреймовая модель;
- ◆ модель семантической сети.

Продукционная модель основывается на множестве правил. Схематически правила в продукционной системе можно представить в обобщенной форме: $P_1, \dots, P_m \rightarrow Q_1, \dots, Q_n$, которая читается следующим образом: если предпосылки P_1 и P_m верны, то можно выполнить действия Q_1 и Q_n . Предпосылки часто называют условиями, а действия – заключениями. Продукционная модель чаще всего применяется в промышленных экспертных системах. Она хороша своей наглядностью, высокой модульностью, простотой механизма логического вывода, легкостью внесения дополнений и изменений.

Логические модели основываются на классическом исчислении предикатов I-го порядка, когда предметная область или задача описывается в виде набора аксиом. Исчисление предикатов I-го порядка в промышленных экспертных системах практически не используется [6].

Фреймовая модель определяется как структура данных для представления стереотипных ситуаций [12]. Эта структура содержит разнообразную информацию: об объектах и событиях, которые следует ожидать в той или иной ситуации, о том, как использовать информацию, имеющуюся во фрейме. Фрейм F является средством, которое может связать декларативные и процедурные знания о некоторой сущности в структуру записей, которая состоит из слотов-*vi* и наполнителей-*qj*. Системы фреймов оказываются значимыми в САПР, так как они создают средства структурирования эвристических знаний, связанных с применением правил и классификацией объектов. При использовании фреймов эвристические знания не «размываются» по программному коду приложения, но и не собираются воедино в виде метазнаний, а распределяются между теми видами объектов, к которым они приложены, и существуют на уровне управления в иерархии представления таких объектов [1].

Семантическая сеть – это ориентированный граф, вершины которого – понятия, а дуги – отношения между ними [6]. В качестве понятий обычно выступают абстрактные или конкретные объекты, а отношения – это связи. Наиболее часто в семантических сетях используют отношения: связи «часть-целое», «класс-подкласс», «элемент-множество» и т.п.; функциональные связи; количественные; пространственные; временные; атрибутивные; логические связи; лингвистические. Недостатком семантической сети является сложность процедуры поиска вывода на семантической сети.

Форма представления знаний существенно влияет на характеристики и свойства системы, основанной на знаниях, поэтому представление знаний является одной из важных проблем в области интеллектуализации САПР и искусственного интеллекта (ИИ).

Отметим, что проблему представления знаний нельзя отделять от проблемы определения состава знаний. Эти две проблемы взаимообусловлены. Действительно, выбранный способ представления может оказаться непригодным в принципе либо неэффективным для выражения некоторых знаний. Вопрос представления знаний можно разделить на две независимые задачи: как организовать (структурировать) знания и как представить знания в выбранной форме. Стремление выделить организацию знаний в самостоятельную задачу вызвано тем, что эта задача возникает для любого языка представления и способы решения этой задачи являются практически схожими, вне зависимости от используемого формализма.

При решении проблемы представления знаний необходимо рассмотреть следующие вопросы:

- ◆ определение состава представляемых знаний;
- ◆ определение организации знаний;
- ◆ определение формы представления знаний.

В частности, для ЭС представление знаний, т.е. определение модели представления знаний, определяется следующими факторами:

- ◆ проблемной средой;
- ◆ архитектурой экспертной системы;
- ◆ потребностями и целями пользователей;
- ◆ языком общения.

При проектировании модели знаний в САПР очень важно учитывать архитектуру экспертной системы. И с учетом архитектуры экспертной системы знания целесообразно делить на интерпретируемые и неинтерпретируемые. К первому типу относятся те знания, которые сам решатель способен интерпретировать (интерпретатор). Все остальные знания относятся ко второму типу. Решатель не знает их структуры и содержания. Если эти знания используются каким-либо компонентом системы, то он не "осознает" этих знаний. Неинтерпретируемые знания подразделяются на вспомогательные знания, хранящие информацию о лексике и грамматике языка общения, о структуре диалога, и поддерживающие знания. Вспомогательные знания обрабатываются естественно-языковой компонентой, но процесс этой обработки решатель не осознает, так как этот этап обработки входных сообщений является вспомогательным для проведения экспертизы. Поддерживающие знания используются при создании системы и при выполнении объяснений. Поддерживающие знания исполняют роль описаний, как интерпретируемых знаний, так и действий системы. Такие знания подразделяются на технологические и семантические знания. Технологические знания содержат сведения о времени создания описываемых ими знаний, об авторе знаний и т.п. Семантические поддерживающие знания содержат смысловое описание этих знаний. Они содержат информацию о причинах ввода знаний, о назначении знаний, описывают способ использования знаний и получаемый эффект. Поддерживающие знания имеют описательный характер.

Интерпретируемые знания можно разделить на предметные знания, управляющие знания и знания о представлении. Знания о представлении содержат информацию о том, каким образом (в каких структурах) в системе представлены интерпретируемые знания.

Предметные знания содержат данные о предметной области и способах преобразования этих данных при решении поставленных задач. Отметим, что по отношению к предметным знаниям знания о представлении и знания об управлении являются метазнаниями. В предметных знаниях можно выделить описатели и собственно предметные знания. Описатели содержат определенную информацию о предметных знаниях такую, как: коэффициент определенности правил и данных, меры важности и сложности. Предметные знания разбиваются на факты и исполняемые утверждения. Факты определяют возможные значения сущностей и характеристик предметной области. Исполняемые утверждения содержат информацию о том, как можно изменять описание предметной области в ходе решения задач. Другими словами, исполняемые утверждения – это знания, задающие процедуры обработки. Эти знания могут быть заданы не только в процедурной, но и в декларативной форме [13].

Качественные и количественные показатели экспертной системы могут быть улучшены за счет использования метазнаний, т.е. знаний о знаниях. Метазнания не представляют некоторую единую сущность, они могут применяться для достижения различных целей. Перечислим возможные назначения метазнаний:

- 1) метазнания в виде стратегических метаправил используются для выбора релевантных правил;
- 2) метазнания используются для обоснования целесообразности применения правил из области экспертизы;
- 3) метаправила используются для обнаружения синтаксических и семантических ошибок в предметных правилах;
- 4) метаправила позволяют системе адаптироваться к окружению путем перестройки предметных правил и функций;

5) метаправила позволяют явно указать возможности и ограничения системы, т.е. определить, что система знает, а что не знает.

Вопросы организации знаний необходимо рассматривать в любом представлении, и их решение в значительной степени не зависит от выбранного способа (модели) представления. Выделим следующие аспекты проблемы организации знаний:

- ◆ организация знаний по уровням представления и по уровням детальности;
- ◆ организация знаний в рабочей памяти;
- ◆ организация знаний в базе знаний.

Для того чтобы экспертная система могла управлять процессом поиска решения, была способна приобретать новые знания и объяснять свои действия, она должна уметь не только использовать свои знания, но и обладать способностью понимать и исследовать знания, т.е. экспертная система должна иметь знания о том, как представлены ее знания о проблемной среде. Если знания о проблемной среде назвать знаниями нулевого уровня представления, то первый уровень представления содержит метазнания, т.е. знания о том, как представлены во внутреннем мире системы знания нулевого уровня. Первый уровень содержит знания о том, какие средства используются для представления знаний нулевого уровня. Знания первого уровня играют существенную роль при управлении процессом решения, при приобретении и объяснении действий системы. В связи с тем, что знания первого уровня не содержат ссылок на знания нулевого уровня, знания первого уровня независимы от проблемной среды. Число уровней представления может быть больше двух. Второй уровень представления содержит сведения о знаниях первого уровня, т.е. знания о представлении базовых понятий первого уровня. Разделение знаний по уровням представления обеспечивает расширение области применимости системы [14].

Выделение уровней детальности позволяет рассматривать знания с различной степенью подробности. Количество уровней детальности во многом определяется спецификой решаемых задач, объемом знаний и способом их представления. Как правило, выделяется не менее трех уровней детальности, отражающих соответственно общую, логическую и физическую организацию знаний. Введение нескольких уровней детальности обеспечивает дополнительную степень гибкости системы, это позволяет производить изменения на одном уровне, не затрагивая другие. Изменения на одном уровне детальности могут приводить к дополнительным изменениям на этом же уровне, что оказывается необходимым для обеспечения согласованности структур данных и программ.

Рассмотрев различные модели и подходы к организации знаний, необходимо отметить, что процесс проектирования требует применения различных моделей знаний, так как различные этапы творческой деятельности конструктора не представляется возможным описать одной моделью. Поэтому при создании экспертной системы необходимо применительно к её архитектуре обеспечить эффективное использование всех имеющихся в системе видов знаний, что возможно при построении её в виде гибридной экспертной системы. Разнородные знания целесообразно классифицировать и для каждого вида использовать определённую модель представления знаний.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Джексон П. Введение в экспертные системы. – М.: Издательский дом «Вильямс», 2001. – 624 с.
2. Дюк В., Самойленко А. Data Mining: учебный курс. – СПб.: Питер, 2001. – 368 с.

3. *Городецкий В.И., Самойлов В.В., Малов О.А.* Современное состояние технологии извлечения знаний из баз и хранилищ данных. Ч. I // *Новости искусственного интеллекта.* – 2002, №3. – С. 3-12.
4. *Осипов Г.С.* Приобретение знаний интеллектуальными системами. – М.: Наука, 1997. – 112 с.
5. *Boose J. and Gaines B.* Knowledge Acquisition Tools for Expert Systems. New York: Academic Press. 1988. p. 620.
6. *Гаврилова Т.А. Хорошевский В.Ф.* Базы знаний интеллектуальных систем. – СПб: Питер, 2000. – 384 с.
7. *Kingston J., Doheny J., Filby I.* Evaluation of workbenches which support the Common KADS methodology. Knowledge Engineering Review. 1995. №10.
8. *Орлов С.А.* Технологии разработки программного обеспечения. – СПб.: Питер, 2002. – 464 с.
9. *Попов Э.В., Фоминых И.Б., Кисель Е.Б., Шапот М.Д.* Статические и динамические экспертные системы // Под ред. Попова Э.В. – М.: Финансы и статистика, 1996. – 320 с.
10. *Поспелов Д.А.* Данные и знания. Искусственный интеллект // В 3 кн. Кн. 1. – М.: Радио и связь, 1990. – 384 с.
11. *Алексеев А.В., Борисов А.Н. и др.* Интеллектуальные системы принятия проектных решений. – Рига: Зинатне, 1997. – 320 с.
12. *Минский М.* Фреймы для представления знаний / Пер. с англ. – М.: Энергия, 1979. – 152 с.
13. *Пупков К.А., Кольков В.Г.* Интеллектуальные системы. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2003. – 348 с.
14. *Курейчик В.М., Лебедев Б.К., Нужнов Е.В.* Применение экспертных систем в инженерной практике. – Таганрог: ТРТУ, 1996. – 135 с.

УДК 518.5:331.108.26

С.В. Скороход

ОТБОР ПЕРСОНАЛА В УСЛОВИЯХ НЕЧЁТКОЙ ИНФОРМАЦИИ НА ОСНОВЕ ЭКСПЕРТНЫХ ОЦЕНОК

Постановка задачи. Проблема оптимального подбора кадров является актуальной для любого предприятия. От её решения зависят качество и отдача каждого сотрудника и конечный успех предприятия в целом.

Характерными особенностями данной проблемы являются субъективность и неопределённость. Субъективность выражается в том, что оценку и подбор персонала производят люди со своими стереотипами, симпатиями, предпочтениями, которые иногда могут идти вразрез с целями самого предприятия. Неопределённость проявляется в неточных формулировках требований типа «уверенное владение Microsoft Word», «опытный пользователь ПК», «умение работать в команде», допускающими трактовку в очень широких пределах.

Одним из наиболее достоверных методов профессионального отбора является выполнение кандидатами специально подготовленных заданий с последующей оценкой результата группой экспертов [1]. Несмотря на преимущества данного метода, его недостатком является слабая формализованность процесса обработки результатов тестирования.

Задачей данной работы является разработка математической модели для сравнения кандидатов по результатам тестирования с использованием теории нечётких множеств и нечёткой математики [2, 3], которые позволяют формализовать присущие этому процессу свойства субъективности и неопределённости.

Общая схема отбора. Общая схема отбора изображена на рис. 1.