

## Применение принципов организации интернет-конференций при разработке экспертных систем

Зеленёва И.Я., Ковалёв С.А., Лаврик А.С.

Донецкий национальный технический университет, г. Донецк  
кафедра электронных вычислительных машин

### **Abstract**

*Zelenyova I.J., Kovalev S.A., Lavrik A.S. Application of Internet-conferences arrangement principles for expert system development. Analysis of contemporary technical Internet-conferences in interpretation for expert systems is offered. On the basis of analysis, method of creation hybrid expert systems is given. These systems allow minimizing or total exclude some fundamental problems apply to classical expert systems.*

### **Введение**

Экспертные системы (ЭС) возникли как значительный практический результат в применении и развитии методов искусственного интеллекта (ИИ) [1].

Область ИИ имеет более чем сорокалетнюю историю развития. С самого начала в ней рассматривался ряд весьма сложных задач, которые, наряду с другими, и до сих пор являются предметом исследований: автоматические доказательства теорем, автоматический перевод с одного естественного языка на другой, распознавание изображений и анализ сцен, планирование действий роботов, алгоритмы и стратегии игр.

Классическая ЭС - это комплекс программ, выполняющий функции эксперта при решении задач из определённой предметной области. ЭС выдают советы, проводят анализ, ставят диагноз. Практическое применение экспертных систем на предприятиях способствует эффективности работы и повышению квалификации специалистов [1].

Главным достоинством экспертных систем является возможность накопления знаний и сохранение их длительное время. В отличие от человека, к любой информации экспертные системы подходят объективно, что улучшает качество проводимой экспертизы. При решении задач, требующих обработки большого объема знаний, возможность возникновения ошибки при переборе очень мала.

При создании ЭС возникает ряд затруднений. Это, прежде всего, связано с тем, что заказчик не всегда может точно сформулировать свои требования к разрабатываемой системе. Также возможно возникновение трудностей чисто психологического порядка: при создании базы знаний

системы эксперт может препятствовать передаче своих знаний, опасаясь, что впоследствии его заменят “машиной”. Кроме того, довольно сложно создать самообучающуюся систему, которая была бы постоянно актуальна и соответствовала новым открытиям в проблемной области [2].

Для решения описанных выше проблем предлагались различные методы построения экспертных систем - от реализации ЭС, подобных доске объявлений (HEARSAY-III, AGE) [2], до создания систем в WWW (W.E.S.T) [3,4], позволяющим иметь доступ к базе знаний большому количеству пользователей, - но проблема постоянной актуализации базы знаний полностью решена не была.

В данной статье проанализирован принцип организации современных технических Интернет-форумов в аспекте применения к экспертным системам, что открывает новые возможности для минимизации либо полного исключения перечисленных проблем. На основе проведенного анализа предлагается методика создания гибридных экспертных систем, называемых в данном контексте гибридными.

### **Основные определения**

Экспертная система, как и эксперт-человек, в процессе своей работы оперирует с необходимым набором знаний в строго определённой предметной области. Вся информация, необходимая для работы ЭС, определенным образом формализована и представлена в памяти ЭВМ в виде базы знаний, которая может изменяться и дополняться в процессе развития системы [1].

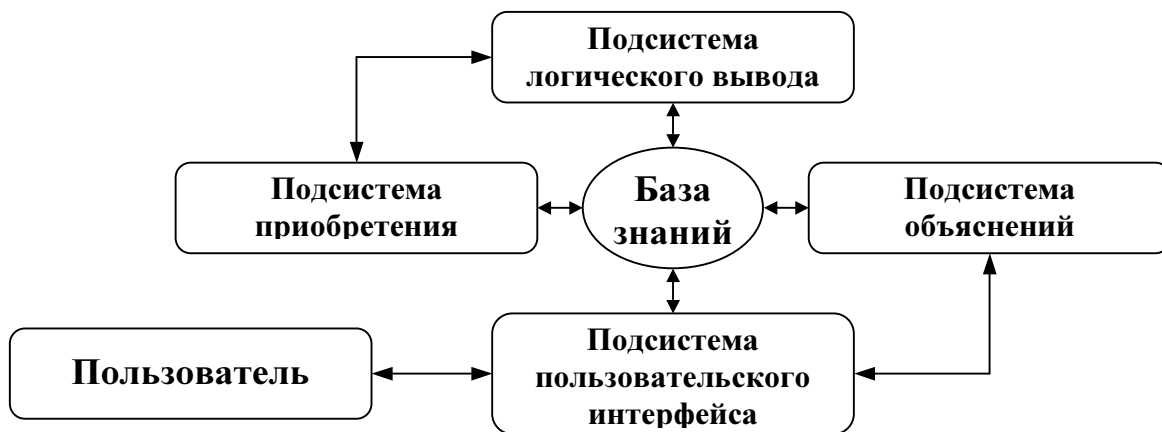


Рисунок 1 – Обобщённая структурная схема экспертной системы.

Экспертная система (рис. 1) может полностью взять на себя функции, выполнение которых обычно требует привлечения опыта человека-специалиста, или играть роль ассистента для специалиста, принимающего решение. Другими словами, система (техническая или социальная), требующая принятия решения, может получить его непосредственно от

программы или через промежуточное звено — человека, который общается с программой. Тот, кто принимает решение, может быть экспертом со своими собственными правами, и в этом случае программа может "оправдать" свое существование, повышая эффективность его работы. Альтернативный вариант — человек, работающий в сотрудничестве с такой программой, может добиться с ее помощью результатов более высокого качества [2].

Таким образом, можно утверждать, что наиболее эффективной системой, была бы та, в которой сочетались бы опыт эксперта, а так же база накопленных знаний относительно решения проблем в интересующей области. Системой подобного плана может выступать Интернет-конференция.

Рассмотрим структуру специализированной конференции, и сравним её со структурой классической ЭС в развёрнутой форме (рис. 2).

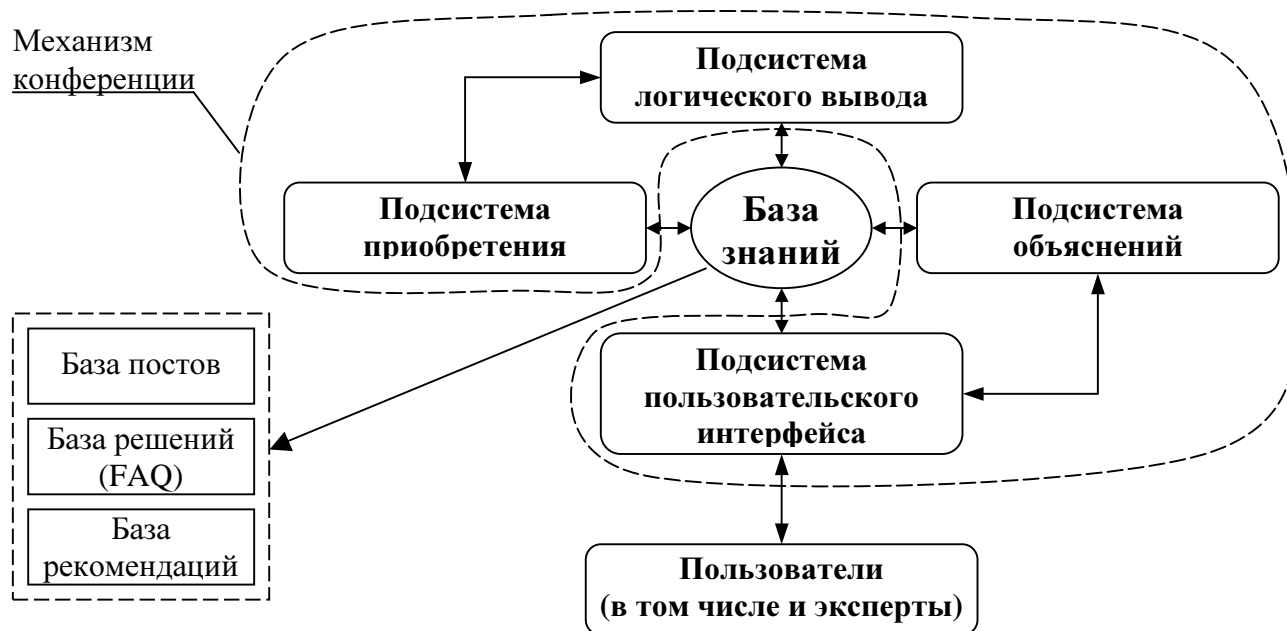


Рисунок 2 – Структура специализированной конференции.

Очевидно, что в основе такой конференции, так же как и в ЭС, лежит база знаний, представленная тремя частями – база постов, база решений, база рекомендаций. **База постов** представляет собой хранилище сообщений, оставленных участниками конференции. Она имеет строго структурированный вид со ступенчатым разбиением на подтемы, аналогично структуре дерева каталогов и согласно классификации объектов проблемной области. **База решений** содержит решения проблем, которые встречаются наиболее часто, для того, чтобы ускорить процесс

помощи пользователю. **База рекомендаций** представляет собой набор статей и заметок с подробными рекомендациями по поводу решения некоторых проблем, она необходима для повышения квалификации пользователя в предметной области.

Всю «обвязку» базы знаний, принимает на себя **механизм конференции**, обеспечивающий как функционирование самой базы, так и связь её с пользователем, которым может быть либо рядовой специалист, столкнувшийся с неизвестной ему проблемой, либо эксперт в данной области.

Каждый пользователь, обращаясь к конференции с проблемой, создаёт сообщение в соответствующем разделе и впоследствии получает ответы на свой вопрос от других участников. Такой интерактивный режим приводит к тому, что решение данной проблемы остаётся в базе и может быть использовано другими пользователями, обратившимися с аналогичным вопросом.

В отличие от классических экспертных систем, Интернет-конференции обладают несколькими функциями, расширяющими возможности ЭС. Такими дополнительными функциями являются: специализированная **подсистема поиска**, позволяющая быстро найти решение проблемы в базе, благодаря сложной системе фильтров, а так же **система личных сообщений**, позволяющая участникам конференции общаться напрямую, что также может ускорить решение возникшей проблемы. При этом недостатком является тот факт, что в дальнейшем другие пользователи не смогут увидеть ранее найденное решение в базе.

Такая организация позволяет решать проблемы классических экспертных систем, представленных в виде одного программного продукта. Благодаря двойной системе поиска ответа – механизм конференции + человек эксперт – значительно ускоряется процесс поиска решения проблемы. База знаний является постоянно актуальной, так как новые решения постоянно добавляются туда с запросами пользователей.

### ***Принцип построения гибридной экспертной системы***

В связи с описанной выше рациональностью создания интерактивных сетевых ЭС, предлагается новый подход к созданию гибридных экспертных систем (ГЭС), сочетающих в себе плюсы классических экспертных систем и интерактивных форумов. В общем случае структура таких систем является двухуровневой (рис. 3).

Одним из уровней системы является множество локальных терминалов (**LT**), каждый из которых интерпретирует классическую ЭС со своей собственной базой знаний (**LB**). Второй уровень представлен глобальным сервером (**Server**), имеющим общую базу знаний (**GB**).

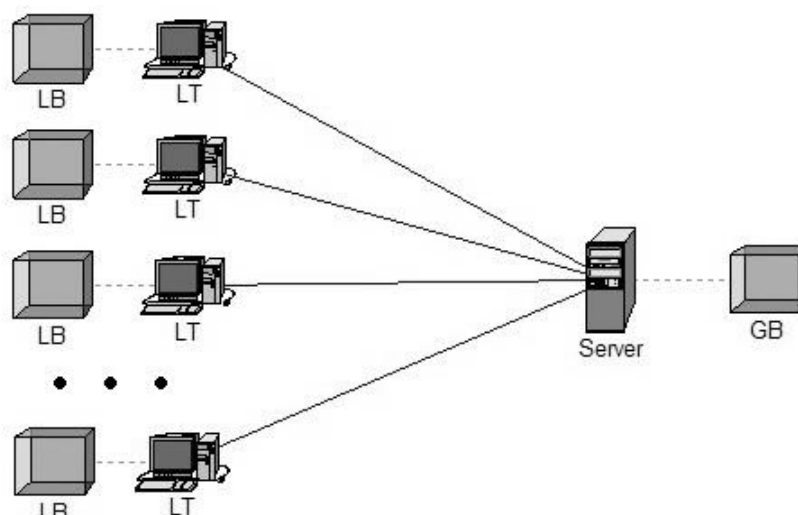


Рисунок 3 – Аппаратная организация двухуровневой гибридной экспертной системы

Разбиение базы на две составляющие обосновано экономией памяти, затрачиваемой на хранение базы знаний. Вариант хранения всей базы на каждом терминале не имеет особого практического значения. Гораздо эффективнее иметь малую базу наиболее часто запрашиваемых решений на каждом терминале и общую базу на сервере, к которой можно обращаться по мере необходимости. Кроме того, такой подход ускоряет поиск решения, так как скорость выборки из базы напрямую зависит от её размеров.

Предлагается следующий *алгоритм работы гибридной ЭС*:

1. Специалист обращается с запросом к локальному терминалу ГЭС.  
 2. Если решение проблемы не найдено, проблемный запрос формализуется и отправляется на глобальный сервер ГЭС с общей базой знаний. Если решение найдено → п.8.

3. Сервер, обрабатывая запрос, просматривает свою базу знаний в поисках решения.

Если решение найдено → п.4, иначе → п.5.

4. Решение отправляется на локальный терминал обратившегося к ГЭС специалиста и остаётся в его локальной базе знаний, → п.7.

5. Запрос пользователя передаётся на терминалы всех, подключенных к ГЭС пользователей, для поиска решения в локальных базах.

Если решение найдено → п.4, иначе → п.6.

6. Запрос оставляется на локальных терминалах на тот случай, если один из специалистов, подключенных к ГЭС, знает решение.

Если решение найдено → п.7, иначе → п.8.

7. Если решений несколько (несколько специалистов разбираются в проблеме), они обрабатываются сервером для получения единого

формализованного вида. Результат обработки записывается в глобальную базу знаний, -> п.8.

#### 8. Формирование ответа ГЭС пользователю.

Такой алгоритм позволяет, с одной стороны, поддерживать актуальными как глобальную, так и локальные базы знаний ГЭС, а с другой стороны – не перегружать сервер, так как наиболее вероятными являются запросы, которые уже имели место на конкретном терминале.

#### **Проблематика дальнейших исследований**

Главными проблемами в предлагаемых ГЭС являются выбор способа распределения базы решений между локальными терминалами и сервером, а так же способ сокращения времени поиска решения.

Рассмотрим проблемы первого типа.

Пусть  $Q_G$  – количество записей в глобальной базе.  $Q_L^i$  – количество записей в локальной базе  $i$ -го терминала ГЭС,  $n$  – количество локальных терминалов в ГС,  $M$  – размер одной записи в базе решений,  $V$  – общий объём дискового пространства ГЭС. Тогда справедливы следующие соотношения:

$$Q_G = \bigcap_{i=1}^n Q_L^i \quad (1)$$

$$Q_L^i \leq Q_G, i = 1..n \quad (2)$$

$$V = M * \sum_{i=1}^n Q_L^i + M * Q_G \quad (3)$$

В худшем случае  $Q_L^i = Q_G$ . Это возможно при очень малых значениях  $n$ . В таком случае каждый локальный терминал хранит полную базу знаний, поэтому экономии дискового пространства нет.

$$V = M * \sum_{i=1}^n Q_L^i + M * Q_G = (Q_G * (n+1)) * M \quad (4)$$

С увеличением  $n$  растёт разница между  $Q_L^i$  и  $Q_G$ . Таким образом, при достаточно больших значениях  $Q_G \gg Q_L^i$  будет справедливо отношение

$\sum_{i=1}^n Q_L^i \approx Q_G$ . В таком случае, объём дискового пространства можно определить, как

$$V = M * \sum_{i=1}^n Q_L^i + M * Q_G = 2 * Q_G * M \quad (5)$$

Проблемы, связанные со временем поиска решения, зависят от двух факторов – времени поиска по базе и времени отправки/получения запросов от сервера и к терминалам.

Время поиска в базе уменьшается за счёт применения более эффективных алгоритмов.

Стандартный поиск перебором в данной ситуации является неэффективным, так как сильно зависит от размеров базы – количество итераций алгоритма, равно количеству записей в базе, поэтому при больших объёмах данных время, затрачиваемое на выборку решения, будет недопустимо велико [5].

Более быстрым алгоритмом является двоичный (бинарный) поиск, требующий упорядоченности базы.

Число операций алгоритма бинарного поиска имеет порядок  $\log_2(n)$ , где  $n$  - число элементов массива. При многочисленных обращениях к этой процедуре число операций будет равно  $m \cdot \log_2(n)$  (где  $m$  - число обращений). Ускорения этой процедуры можно добиться за счет сохранения предыдущего результата операции и попыток поиска при новом обращении в ближайших узлах массива с дальнейшим расширением области поиска в случае неуспеха [5].

При этом в наихудших случаях число операций будет больше (примерно в 2 раза) в сравнении с бинарным поиском, но обычно при  $m$  значительно большем, чем  $\log_2(n)$ , удается довести порядок числа операций до  $m$ , то есть сделать его почти независимым от размера массива.

Таким образом, необходимо выделить на сервере небольшие области памяти под каждый локальный терминал для хранения истории предыдущих запросов. Но это потребует лишнего дискового пространства.

Так же можно использовать алгоритм интерполяционного поиска, который наилучшим образом проявляет себя при больших объёмах просматриваемых данных.

Пусть  $K$  - это искомый ключ в базе,  $K_l, K_u$  - ключи, определяющие границы поиска элемента (изначально границы базы),  $l, u$  - индексные номера ключей, соответственно.

Если известно, что  $K$  лежит между  $K_l$  и  $K_u$ , то следующая проба делается на расстоянии  $(u-l)(K-K_l)/(K_u-K_l)$  от  $l$ , предполагая, что ключи являются числами, возрастающими приблизительно в арифметической прогрессии.

Интерполяционный поиск работает за  $\log(\log(n))$  операций, если данные распределены равномерно [6].

Проблемы задержек трафика между сервером и терминалами зачастую решаются улучшением аппаратного обеспечения сети – заменой среды передачи данных на более быструю (например, замена Ethernet на Fast Ethernet); улучшением материальной базы серверов, что приводит к ускорению обработки скриптов; установкой современных роутеров, ускоряющих обмен пакетами между станциями сети и т.д.

## **Заключение**

В связи с развитием аппаратных и программных средств, стало возможным построение гибридных интерактивных экспертных систем. Предложенный авторами алгоритм работы такой системы, а так же принцип раздельного хранения информации позволяют достичь экономии дискового пространства и повысить вероятность нахождения решения в короткие сроки. Дальнейшие исследования авторов в этой области связаны разработкой и исследованием способов распределения информации в ГЭС, а так же поддержка актуальности баз знаний. Ожидается, что наибольшую эффективность такие системы могут обеспечивать при наличии большого числа специалистов (локальных баз), подключенных к глобальному серверу.

## **Литература**

1. Джексон. П., Введение в экспертные системы. – К.: Диалектика, 2001. – 624 с.
2. Хейес-Рот Ф., Уотерман Д., Ленат Д., Построение экспертных систем. – М.: Мир, 1987. – 430 с.
3. Коптев В.Ю., Савченко О.Ф., Вольф В.Т., Боброва Т.Н., Гребенникова И.Г. Особенности построения информационно-диагностических систем в эпизоотологии. – Материалы междунар. Конф. «ИСТ-2003», Новосибирск: НГТУ, 2003. – Т.3, с.43.
4. Гаврилов А.В., Новицкая Ю.В. Инструментальное программное обеспечение для создания гибридных экспертных систем. – Труды Межд. Конф. ИСТ-2000, Новосибирск, НГТУ, 2000, т.3. – С. 488-490.
5. Kurtz S., Fundamental Algorithms for a Declarative Pattern Matching System: Bisher erschienene Reports an der Technischen Fakultät., Universität Beilfield., 1995. – 247 с.
6. Кнут Д., Искусство программирования: “Сортировка и поиск”: Монография. – М.: Мир, 1978. – 675 с.