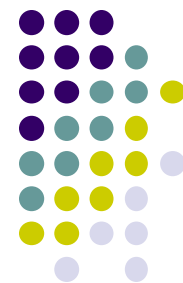




Системы управления знаниями

Методы работы со знаниями в ЭС

Учебные вопросы.



- 1. Методы извлечения знаний от экспертов.
- 2. Формализация качественных знаний.



- ***Приобретением знаний*** называется выявление знаний из источников и преобразование их в нужную форму, а также перенос в базу знаний ЭС.
- Источниками знаний могут быть книги, архивные документы, содержимое других баз знаний и т.п. (*объективизированные знания*).
- Другим типом знаний являются *экспертные знания*, которые имеются у специалистов, но не зафиксированы во внешних по отношению к нему хранилищах.



- **Экспертные знания являются субъективными.**
- Еще одним видом субъективных знаний являются *эмпирические знания*. Такие знания могут добываться ЭС путем наблюдения за окружающей средой (если у системы есть средства наблюдения).

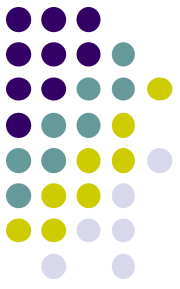


- Ввод в базу знаний объективизированных знаний не представляет особой проблемы, тогда как *выявление и ввод экспертных знаний является достаточно сложной задачей.*

1. Методы извлечения знаний от экспертов



- Возможны три режима взаимодействия инженера по знаниям с экспертом-специалистом:
- *протокольный анализ,*
- *интервью,*
- *игровая имитация профессиональной деятельности.*



- Протокольный анализ заключается в фиксации «мыслей вслух» эксперта во время решения проблемы и в последующем анализе полученной информации.
- В режиме интервью инженер по знаниям ведет с экспертом активный диалог, направляя его в нужную сторону.



- При игровой имитации эксперт помещается в ситуации, похожие на те, в которых протекает его профессиональная деятельность.
- Наблюдая за его действиями в различных ситуациях, инженер по знаниям формирует свои соображения об экспертных знаниях, которые впоследствии могут быть уточнены с экспертом в режиме интервью.



- Каждый из способов извлечения знаний имеет *свои преимущества и недостатки*.
- Так, при анализе протоколов инженеру по знаниям **нелегко отделить понятия**, **важные** для включения в словарь предметной области, от тех, которые при «мыслях вслух» появляются случайно.
- Кроме того, в протоколах обнаруживаются **«пробелы»**, когда рассуждение эксперта как бы прерывается и продолжается уже на основе пропущенных шагов вывода.



- Заполнение подобных «пробелов» возможно лишь в режиме интервью.
- Таким образом, во всех трех подходах к извлечению знаний из экспертов необходим этап интервью, что делает его одним из важнейших методов приобретения знаний.



- Существует достаточно много стратегий интервьюирования.
- Наиболее известны три:
- ***разбиение на ступени***,
- ***репертуарная решетка***,
- ***подтверждение сходства***.
- При *разбиении на ступени* эксперту предлагается назвать наиболее важные, по его мнению, понятия предметной области и указать между ними отношения структуризации, т.е. отношения типа «род-вид», «элемент-класс», «целое-часть» и т.п.



- Например, относительно двух событий: «рост заработной платы» и «повышение уровня жизни» эксперт сообщил, что «рост заработной платы» обычно сопровождается «повышением уровня жизни». Тогда возникают вопросы:
- а) повышение уровня жизни **всегда** сопровождается ростом заработной платы?
- б) повышение уровня жизни **может** сопровождаться ростом заработной платы?



- Ответ эксперта а) будет свидетельствовать о том, что исходное сообщение относится к типу «***a*** есть проявление ***b***, и ***b*** сопровождается ***a***» ;
- Ответ б) - к типу «***a*** сопровождается ***b***, и ***b*** может сопровождаться ***a***» .



- С каждым типом сообщения связывается формальная конструкция базы знаний ЭС, то есть бинарное отношение на множестве объектов (событий) или конъюнкция отношений:

<имя объекта> <имя отношения> <имя объекта>
<**a**><проявление><**b**>И<**b**><сопровождается><**a**>



- *Стратегия репертуарной решетки* направлена на выявление характеристических свойств понятий, позволяющих отделять одни понятия от других.
- Методика состоит в предъявлении эксперту троек понятий с предложением назвать признаки для каждой двух понятий, которые отделяли бы их от третьего.



- Так как каждое понятие входит в несколько троек, то на основании такой процедуры происходит уточнение объемов понятий и формируются «симптокомплексы» понятий, с помощью которых эти понятия могут идентифицироваться в базе знаний.



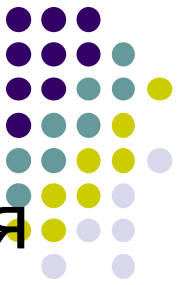
- Например, эксперту в области представления знаний предъявляется триада понятий, описывающих способы представления: **семантические сети, фреймы, системы продукций.**
- Эксперту предлагается ответить на следующие вопросы:
- *Какой из указанных способов представления отличается от двух других?* **системы продукций;**



- *Какое свойство отличает системы
продукций от семантических сетей и
фреймов? **легкость описания динамики;***
- *Назовите противоположное свойство
свойству «легкость описания динамики»
трудность описания динамики;*
- *Дайте имя свойству, имеющему
значения «легкость описания динамики»
и трудность описания динамики»
ВОЗМОЖНОСТЬ описания динамики.*



- **В результате формируется шкала** с именем «возможность описания динамики» и значением «легкость описания динамики» для объекта «системы продукций»; «трудность описания динамики» для объектов «семантические сети» и «фреймы».
- Предлагая эксперту аналогичные вопросы об отличии семантических сетей от систем продукций и фреймов, можно выявить и другие свойства понятий.

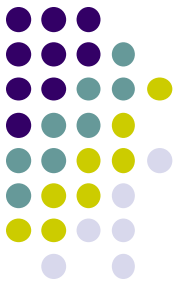


- *Стратегия подтверждения сходства* состоит в том, что эксперту предлагается установить принадлежность каждой пары понятий из предметной области к некоторому отношению сходства (толерантности).
- Для этого эксперту задается последовательность достаточно простых вопросов, цель которых заключается в уточнении того понимания сходства, которое вкладывает эксперт в утверждение о сходстве двух понятий предметной области.

2. Формализация качественных знаний



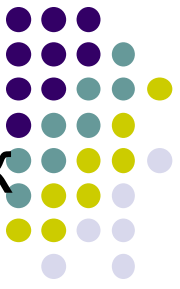
- При работе с экспертами на естественном языке возникает **проблема формализации качественных знаний**, которая может быть решена с использованием теории нечетких множеств.



- Одним из основных понятий теории НМ является понятие *лингвистической переменной*.
- Значениями лингвистических переменных являются не числа, а слова или предложения некоторого искусственного либо естественного языка.



- Например, числовая переменная «возраст» принимает дискретные значения между нулем и сотней, а целое число является значением переменной.
- Лингвистическая переменная «возраст» может принимать значения: «молодой», «старый», «довольно старый», «очень молодой» и т.д.
- Эти термины - *лингвистические значения переменной*. Множество допустимых значений лингвистической переменной называется *терм-множеством*.



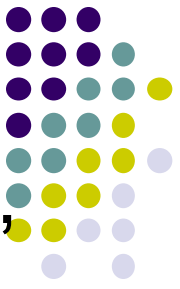
- Для перехода от качественных описаний к формализованным необходимо **построить функции принадлежности.**
- Можно выделить две группы методов построения функций принадлежности: *прямые* и *косвенные*.
- В *прямых методах* эксперт непосредственно задает правила определения значений функции принадлежности (таблицей, формулой или примером).

- В *косвенных методах* значения функции принадлежности выбираются таким образом, чтобы удовлетворялись заранее сформулированные условия.
- Экспертная информация является только исходной для дальнейшей обработки.
- Примерами дополнительных условий могут служить следующие: функция принадлежности должна отражать близость к заранее выделенному эталону, результатом процедуры обработки должна быть функция принадлежности, удовлетворяющая условиям интервальной шкалы и т.п.





- Рассмотрим процедуру формализации нечеткого понятия «небольшая концентрация барбитуратов в крови».
- Согласно мнениям экспертов *носителем нечеткого множества* формализуемого понятия является конечное множество:
 $N_S = \{0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 1,0\},$
- каждый элемент которого представляет собой определенную концентрацию барбитуратов в крови в миллиграмм-процентах.



- Экспертами формируется матрица оценок, отражающих их мнение о степени выраженности у элементов носителя нечеткого множества N_S формализуемого свойства.
- Степени принадлежности элементов нечеткому множеству определяется посредством парных сравнений, при этом используются следующие оценки.



Таблица 1. Шкала относительной важности при парном сравнении

| Степень предпочтения | Определение | Объяснение |
|----------------------|--|-----------------------------------|
| 0 | Независимы | Эксперт затрудняется ответить |
| 1 | Равная важность (значимость) | Имеют одинаковую важность |
| 2 | Некоторое (слабое) предпочтение значимости | Разница в важности равна 1 баллу |
| 3 | Сильное предпочтение значимости | Разница в важности равна 2 баллам |
| 4 | Очень сильное предпочтение значимости | Разница в важности равна 3 баллам |
| 5 | Абсолютное предпочтение значимости | Разница в важности равна 4 баллам |



- С помощью метода парных сравнений были получены следующие значения функции принадлежности:

$$\mu_S(x) = \{0,1; 0,3; 0,8; 0,9; 1,0; 1,0; 0,9; 0,8; 0,3; 0,1\}.$$

- Таким образом, **нечеткое множество S** имеет вид:

$$S = \{0,1/0,1; 0,3/0,2; 0,8/0,3; 0,9/0,4; \mathbf{1,0/0,5}; \mathbf{1,0/0,6}; 0,9/0,7; 0,8/0,8; 0,3/0,9; 0,1/1,0\}.$$

Отсюда следует, что формализуемому нечеткому понятию полностью соответствует концентрация от 0,5 до 0,6 миллиграмм-процентов, в меньшей степени – концентрация от 0,1 до 0,4 и от 0,7 до 1,0 миллиграмм-процентов; концентрация меньше 0,1 и больше 1,0 миллиграмм-процентов понятием «небольшая» охарактеризована быть не может.

