



Информационный процесс представления знаний

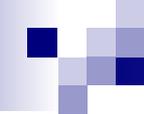
- 
- 1 Интеллектуальные информационные системы.**
 - 2 Свойства и типы знаний.**
 - 3 Модели представления знаний.**
 - 4 Приобретение и формализация знаний.**

1. Интеллектуальные информационные системы.

Одним из основных путей повышения качества управления сложными организационными системами является создание *интеллектуальных информационных технологий (ИИТ)*. Оно связано с решением комплекса проблем синтеза *базы знаний (БЗ)* в *экспертных системах (ЭС)*.



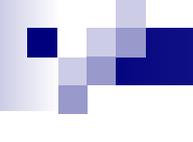
Экспертная система является
средством информационной
технологии, автоматизирующим
процесс представления знаний и его
процедур - получения и генерации
(вывода) знаний.



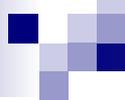
Создание и модификация базы знаний осуществляется совместными усилиями *эксперта (Э)* и *инженера по знаниям (ИЗ)*. Для этой цели создается *интеллектуальный редактор БЗ*, представляющий собой программу диалогового взаимодействия, облегчающую работу с базой знаний.



Решатель (блок логического вывода) производит вывод (генерацию) нового знания, то есть решает поставленную задачу, на основе имеющихся в базе знаний.

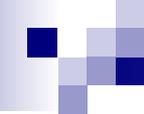


При желании пользователь ЭС может получить объяснение того, как была решена задача. Для этого в ЭС включают блок объяснений. Взаимодействие с экспертной системой пользователя происходит при помощи *интерфейса пользователя*. Центральным блоком экспертной системы является ***база знаний***.



Функции интеллектуальной информационной технологии:

- 1) описывать знания с помощью языков представления знаний;
- 2) организовывать накопление, хранение, анализ, обобщение и структурирование знаний;
- 3) вводить новые знания и объединять их с существующими в СИИ;

- 
- 4) выводить новые знания из имеющихся, оперировать с неполными и неточными знаниями;
 - 5) устранять устаревшие знания, быстро находить требуемые, проверять непротиворечивость накопленных знаний;
 - 6) осуществлять интеллектуальный интерфейс между пользователем и знаниями.

2. Свойства и типы знаний.

Знания - это особая форма информации, представляющая собой совокупность структурированных теоретических и эмпирических положений предметной области, которые представлены в различной форме, обладают определенными свойствами и связаны синтаксическими, семантическими и прагматическими отношениями и которые позволяют решать

Свойства знаний:

1. *Внутренняя интерпретируемость.*

В памяти ЭВМ хранятся не только отдельные информационные единицы, но и системы имен, связанные с ними. Наличие системы имен позволяет системе „знать“, что храниться в ее памяти, и, следовательно, уметь отвечать на запросы о содержании памяти.

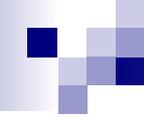
2. Рекурсивная структурируемость.

Информационные единицы могут при необходимости расчленяться на более мелкие и объединяться в более крупные по принципу матрешки. Для этих операций могут использоваться родовидовые отношения и принадлежности элементов к классу.



3. *Взаимосвязь единиц.* Между единицами возможно установление самых разнообразных отношений, отражающих семантику и прагматику связей явлений и фактов.

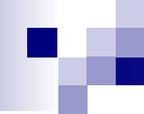
4. Возникновение семантического пространства с метрикой, характеризующего близость-удаленность информационных единиц. Специалисты считают, что знания не могут быть бессистемным „сборищем“ отдельных информационных единиц, а должны быть взаимосвязанными и взаимозависимыми.



5. *Активность.* Активность базы знаний позволяет СИИ формировать мотивы, ставить цели и строить процедуры для их выполнения.

Классификация знаний.

В зависимости от уровня осмысления различают *декларативные* и *процедурные* знания, *глубинные* и *поверхностные*, *жесткие* и *мягкие* знания, *теоретические* и *эмпирические* знания, *концептуальные* и *экспертные*, *синтаксические*, *семантические* и *прагматические* знания.



Глубинные знания образуются как результат обобщения первичных понятий в некоторые абстрактные структуры, которые могут и не иметь вербального описания.

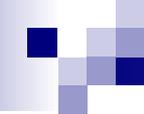
Мягкие знания допускают множественные, расплывчатые решения и приводят к различным вариантам рекомендаций.

Поверхностные знания представляют собой совокупность эмпирических ассоциаций и отношений между понятиями предметной области для стандартных рассуждений и ситуаций.

Концептуальные знания выражают свойства объектов, процессов и ситуаций через понятия (базовые элементы) соответствующей области.

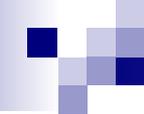
Экспертные знания - это знания специалистов предметной области. Они аккумулируют накопленный практический опыт, навыки и приемы в соответствующей области.

Синтаксические знания характеризуют синтаксическую структуру описываемого объекта или процесса, которая не зависит от смысла и содержания используемых при этом понятий.



Семантические знания содержат информацию непосредственно связанную со знанием и смыслом описываемых объектов и процессов.

Прагматические знания описывают объекты и процессы с точки зрения целей решаемой задачи.



Предметная область – это совокупность элементов, объектов, явлений, процессов, их количественных и качественных характеристик, а также связей между ними, объединенных общей идеей, определенным смыслом или понятием более высокого уровня.

3. Модели представления знаний.

Под представлением знаний подразумевают соглашение о том, как описывать реальную предметную область (понятия и отношения).

Проблемы представления знаний в компьютерных системах решаются на трех уровнях:

 *технический уровень;*

 *-программный (логический) уровень;*

- концептуальный уровень - выработка концепций, моделей, образующих методологию искусственного интеллекта.



В системах искусственного интеллекта используются четыре типа моделей представления знаний : **логические, продукционные, семантические сети и фреймы.**



Логические схемы представляют знания в виде формул, которые состоят из констант, переменных, функций, логических связей и т.д.

В основе всех логических схем представления знаний лежит понятие формальной системы, задаваемой четверкой:

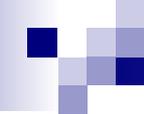
$$M = \langle T, P, A, F \rangle,$$

где **T** - множество базовых элементов (алфавит формальной системы);

P - множество синтаксических правил, позволяющих строить синтаксически правильные выражения **A** из **T**;

A - множество априорно истинных аксиом (любое множество синтаксически правильных выражений);

F - правила вывода, позволяющие расширять множество аксиом.



Преимущества: высокий уровень модульности знаний; лаконичность представления, наличие тела анализа и определение понятие логического вывода, позволяющих получить формальным путем новые знания.

Недостатки: чрезмерный уровень формализации знаний; слабая наглядность, трудность прочтения логических формул и сложность их понимания; низкая производительность систем при отработке знаний и большая требуемая память; отсутствие выразительных средств для отражения особенностей предметной области и структурирования знаний; громоздкость при описании больших объемов знаний

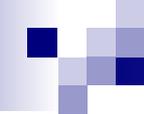
Продукционные модели.

Продукционные правила задаются в виде выражений:

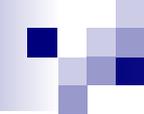
ЕСЛИ условие ТО действие (1)

ЕСЛИ причина ТО следствие (2)

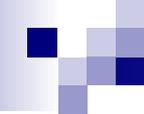
ЕСЛИ ситуации ТО решение (3)



Преимущества: благодаря свойству модульности в продукционные модели можно добавлять и изменять знания (правила, факты); благодаря причинно-следственному характеру правил хорошо отражают прагматическую составляющую знаний.



Недостатки: модель предназначена только для решения небольших задач, при увеличении объема знаний эффективность решения падает.

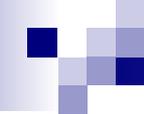


Семантические сети основываются на результатах изучения организации долговременной памяти человека. Для образования своей структуры используют *два компонента* - понятия (объекты, события, процессы, явления) и отношения между понятиями.

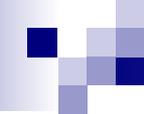
□ *Характеристики сетей:*

- описание объектов производится на естественном языке;
- все знания накапливаются в относительно однородной структуре памяти;
- на сетях определяются унифицированные отношения между объектами, которым соответствуют унифицированные методы вывода;

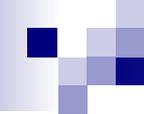
- методы вывода в соответствии с запросами определяют участки семантического знания, имеющего отношение к поставленной задаче, формулируя акт понимания запроса и некоторую цепь выводов, соответствующих решению задачи.



Достоинства: повышенная гибкость; гармоничное и естественное сочетание декларативного и процедурного, синтаксического и семантического знания; наглядность отображения объектов, связей, отношений; легкая читаемость и понимаемость знаний; высокая степень структуризации знаний.



Недостатки: сложность и трудность разработки алгоритмов и их анализа; пассивность структуры сети; требуется разнообразие типов вершин и связей, процедур обработки; трудность представления и обработки неточных и противоречивых знаний.



Фреймы - это особые познавательные структуры, дающие целостное представление о явлениях и их типах. Фреймы отображают концептуальную основу организации памяти человека.

Структура фрейма имеет вид:

$$I: \{ \langle V_1, q_1, p_1 \rangle, \langle V_2, q_2, p_2 \rangle, \dots, \langle V_k, q_k, p_k \rangle \},$$

где I - имя фрейма;

$\langle V_k, q_k, p_k \rangle$ - слот;

V_k - имя слота;

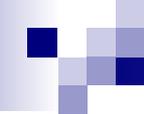
q_k - значение слота;

p_k процедура (является
необязательным элементом).

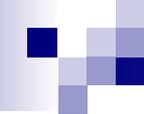


Слот - это структурный элемент фрейма.

Значениями слота могут быть конкретные данные, процедуры и даже продукции. Слот может быть пустым (незаполненным).



Достоинства: высокая
структурируемость, внутренняя
интерпретируемость и связность
слотов и их значений; высокая
наглядность и модульность.



Недостатки: отсутствие
универсальных процедуры их
обработки, что приводит к
неэффективному использованию
ресурсов вычислительной техники
(памяти и быстродействия).

4. Приобретение и формализация знаний.

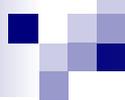
Методы приобретения знаний:

1) Извлечение знаний - процесс приобретения материализованных знаний из источников информации с помощью некоторой совокупности методов и процедур, позволяющих переходить от знаний в текстовой форме к их аналогам для ввода в базу знаний СИИ.



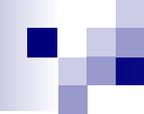
Методы извлечения знаний состоят из текстологических методов и методов автоматической обработки текстов.

Текстологические методы предназначены для получения знаний из материализованных источников: монографий, учебников, статей, методик, инструкций и других носителей профессиональных знаний.



Существенное развитие получили методы извлечения знаний при применении современных информационных технологий, в частности гипертекстовой технологии.

Гипертекст - это организация нелинейной последовательности записи и чтения информации, объединенной на основе ассоциативной связи.



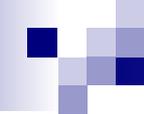
2) Получение знаний - это процесс приобретения знаний эксперта, основанный на использовании непосредственно им самим или инженером по знаниям приемов, процедур, методов и инструментальных средств.



Основной метод получения знаний – коммуникативный. Он основан на прямом диалоге экспертов и инженеров по знаниям как без использования СИИ, так и с применением их (с использованием технологии окон, меню).

Особенности коммуникативных методов:

1. Не имеют формального определения и носят качественный характер.
2. Требуют словесного выражения экспертом своих знаний, что является непростой задачей.
3. Сложность выражения процедурных знаний при их словесном описании.



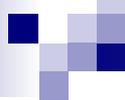
4. Крайняя сложность явного описания знаний, которые являются результатом автоматизма процессов мышления, а также интуиции эксперта.

5. Трудоемкость организации и неэффективность взаимодействия инженера по знаниям и эксперта.

3) Формирование знаний - это процесс автоматического приобретения системой искусственного интеллекта или инструментальным средством нового и полезного знания из исходной и текущей информации, которое в явном виде не формируют эксперты, с целью освоения новых процедур решения прикладных задач на основе использования различных моделей машинного обучения.

4) Приобретение знаний - процесс, основанный на переносе знаний из различных источников в базу знаний путем использования различных методов, моделей, алгоритмов и инструментальных средств.

5) Обучение базы знаний - это процесс ввода (переноса) приобретенных знаний в СИИ на основе применения совокупности методов, приемов и процедур с целью ее заполнения, расширения и модификации.



Основным направлением повышения эффективности процесса представления знаний является его автоматизация.

Выделяют следующие средства автоматизации, которые получили наименование инструментальных средств: языки программирования, языки символьной обработки, языки инженерии (представления) знаний, средства автоматизации проектирования ЭС (интегрированные гибридные среды или комплексы) и оболочки ЭС (пустые или инструментальные ЭС).