

Интеллектуальные информационные системы. Лекция 1.

Основные понятия и определения

Определение интеллектуальной информационной системы

Интеллектуальная информационная система (ИИС) - комплекс программных, лингвистических и логико-математических средств для реализации основной задачи – осуществления поддержки деятельности человека и поиска информации в режиме продвинутого диалога на естественном языке. ИИС являются разновидностью интеллектуальной системы, а также одним из видов информационных систем.

Информационная система - это

совокупность данных (или баз данных), систем управления базами данных и прикладных программ, функционирующих на вычислительных средствах как единое целое для решения определенных задач.

Информационная система является системой обработки данных, хранящихся в базах данных, для манипулирования данными используется язык запросов.

Интеллектуальная система - это

Интеллектуальная система (ИнтС,
англ. *intelligent system*) — это техническая
или программная система, способная решать
задачи, традиционно считающиеся
творческими, принадлежащие конкретной
предметной области, знания о которой
хранятся в памяти такой системы.

Структура интеллектуальной системы

включает три основных блока — базу знаний, механизм вывода решений и интерфейс на естественном языке.

Интеллектуальная система является системой обработки знаний. Для манипулирования знаниями используются языки представления и извлечения знаний.

Отличия между понятиям «данные» и «знания».

Данные – это отдельные факты, характеризующие объекты, процессы и явления предметной области, а так же их свойства.

Знания – это закономерности предметной области (принципы, связи, законы), полученные в результате практической деятельности и профессионального опыта, позволяющие специалистам ставить и решать задачи в этой области.

Этапы обработки данных

Данные как результат измерений и наблюдений



**Данные на материальных носителях информации
(таблицы, протоколы, справочники)**



Модели данных в виде диаграмм, графиков, функций.



Данные в компьютере на языке описания данных



Базы данных на машинных носителях информации

Этапы обработки знаний

Знания в памяти человека как результат мышления



Материальные носители знаний (учебники, методические пособия)



Поле знаний – условное описание основных объектов ПрО, их атрибутов, и отношений между ними



Знания, описанные на языке представления знаний



Базы знаний на машинных носителях информации

Структура интеллектуальной системы включает три основных блока — базу знаний, механизм вывода решений и интерфейс на естественном языке.

Интеллектуальная система является системой обработки знаний. Для манипулирования знаниями используются языки представления и извлечения знаний.

Интеллектуальные информационные системы.

Модели представления знаний

База знаний

База знаний (БЗ; англ. knowledge base, KB) — это особого рода база данных, разработанная для оперирования знаниями (метаданными). База знаний содержит структурированную информацию, покрывающую некоторую область знаний, для использования вычислительной системой (или человеком) с конкретной целью. Современные базы знаний создаются в соответствии с некоторой выбранной моделью представления знаний.

Модели представления знаний

Существуют десятки моделей (или языков) представления знаний для различных предметных областей. Большинство из них можно свести к следующим классам:

- продукционные модели;
- семантические сети;
- фреймы;
- формальные логические модели.

Продукционная модель представления знаний

Продукционная модель или модель, основанная на правилах, позволяет представить знания в виде предложений типа: Если (условие), то (действие).

Под условием (антецедентом) понимается некоторое предложение-образец, по которому осуществляется поиск в базе знаний, а под действием (консеквентом) – действия, выполняемые при успешном исходе поиска.

Продукционная модель представления знаний

Чаще всего вывод на такой базе знаний бывает прямой (от данных к поиску цели) или обратный (от цели для ее подтверждения - к данным). Данные - это исходные факты, хранящиеся в базе фактов, на основании которых запускается машина вывода или интерпретатор правил, перебирающий правила из продукционной базы знаний.

Использование продукционных баз знаний

Продукционная модель чаще всего применяется в промышленных экспертных системах. Она привлекает разработчиков своей наглядностью, высокой модульностью, легкостью внесения дополнений и изменений и простотой механизма логического вывода.

Использование продукционных баз знаний

Имеется большое число программных средств, реализующих продукционный подход (язык OPS 5; "оболочки" или "пустые" ЭС - EXSYS Professional, Карра, ЭКСПЕРТ; ЭКО, инструментальные системы ПИЭС и СПЭИС и др.), а также промышленных ЭС на его основе (например, ЭС, созданных средствами G2) и другие.

Продукционный язык OPS

Являясь универсальным языком программирования, язык OPS в первую очередь предназначен для разработки интеллектуальных систем.

Идеология языка OPS нашла отражение в целом ряде практических реализаций, достаточно сильно отличающихся друг от друга. Одной из первых и наиболее известной является реализация OPS-5, написанная на одной из версий языка Лиспа (Franz LISP). Поэтому синтаксис OPS –5 максимально приближен к синтаксису Лиспа.

Использование версии языка OPS-5

На языке OPS – 5 создан ряд промышленно эксплуатируемых экспертных систем для фирмы DEC с объемом баз знаний от 1000 до 5000 правил. Одной из последних, но уже достаточно широко известной реализацией является OPS-83.

Особенности этой реализации – наличие некоторых конструкций, характерных для процедурных языков программирования, а также сильная типизация данных.

Отличительные черты семейства языков OPS

- ♦ Наличие программного управления стратегий вывода решений;
- ♦ Наличие развитой структуры данных и принципиальной эффективности реализации.

Отличительные черты семейства языков OPS

Язык OPS имеет типичную для продукционных систем архитектуру, включающую в себя базу правил, рабочую память и механизм вывода. База правил состоит из неупорядоченной совокупности правил, рабочая память – из дискретных объектов, называемых элементами рабочей памяти. Элемент рабочей памяти может быть добавлен в рабочую память, удален из нее или модифицирован.

Отличительные черты семейства языков OPS

Механизм вывода является стандартным для системы продукций циклом управления. На первой фазе цикла выбираются все правила, левые части которых сопоставились с содержимым рабочей памяти. На второй фазе правило выполняется.

Встроенный в OPS механизм вывода непосредственно поддерживает только прямой вывод, однако в языке имеются средства для организации обратного и смешанного выводов.

Оболочка CLIPS— среда разработки продукционных баз знаний

CLIPS используется оригинальный LISP-подобный язык программирования, ориентированный на разработку интеллектуальных систем, базирующихся на продукционной модели представления знаний.

Использование языка C в качестве языка реализации объясняется тем, что компилятор LISP не поддерживается частью распространенных платформ.

Оболочка CLIPS

Название CLIPS – аббревиатура от C Language Integrated Production System. Язык был разработан в центре космических исследований NASA (NASA's Johnson Space Center). Первая версия системы вышла в 1984 году, текущая версия 6.3.

Особенности CLIPS

CLIPS является одним из распространенных инструментальных средств разработки интеллектуальных систем. Представляя собой логически полную среду, содержащую встроенный редактор и средства отладки, CLIPS является оболочкой ЭС. Первая версия представляла собой, по сути, интерпретатор порождающих правил. Процедурный язык и объектно-ориентированное расширение CLIPS Object Oriented Language (COOL) были включены в этот программный продукт только в 1990-х годах.

Продукционная модель в среде CLIPS

CLIPS использует продукционную модель представления знаний и включает в себя язык представления порождающих правил и язык описания процедур.

Основными компонентами языка описания правил являются база фактов (fact base) и база правил (rule base). На них возлагаются следующие функции:

- база фактов представляет собой исходное состояние проблемы;
- база правил содержит операторы, которые преобразуют состояние проблемы, приводя его к решению.

Семантические сети

Термин *семантическая* означает "смысловая", а сама семантика - это наука, устанавливающая отношения между символами и объектами, которые они обозначают, то есть наука, определяющая смысл знаков. *Семантическая сеть* - это ориентированный граф, вершины которого - понятия, а дуги - отношения между ними.

Семантические сети

В качестве понятий обычно выступают абстрактные или конкретные объекты, а *отношения* - это связи типа: "это" ("АКО - A-Kind-Of", "is"), "имеет частью" ("has part"), "принадлежит", «учится".

Семантические сети

Характерной особенностью семантических сетей является обязательное наличие трех типов отношений:

класс - элемент класса (цветок - роза);

свойство - значение (цвет - желтый);

пример элемента класса (роза - чайная).

Классификация семантических сетей

По количеству типов отношений:

- Однородные (с единственным типом отношений).
- Неоднородные (с различными типами отношений).

Классификация семантических сетей

По типам отношений:

- Бинарные (в которых отношения связывают два объекта).
- N-арные (в которых есть специальные отношения, связывающие более двух понятий).

Классификация отношений в семантических сетях

Наиболее часто в семантических сетях используются следующие отношения:

- связи типа *"часть - целое"* (*"класс - подкласс"*, *"элемент -множество"*, и т. п.);
- функциональные связи (определяемые обычно глаголами *"производит"*, *"влияет"*...);

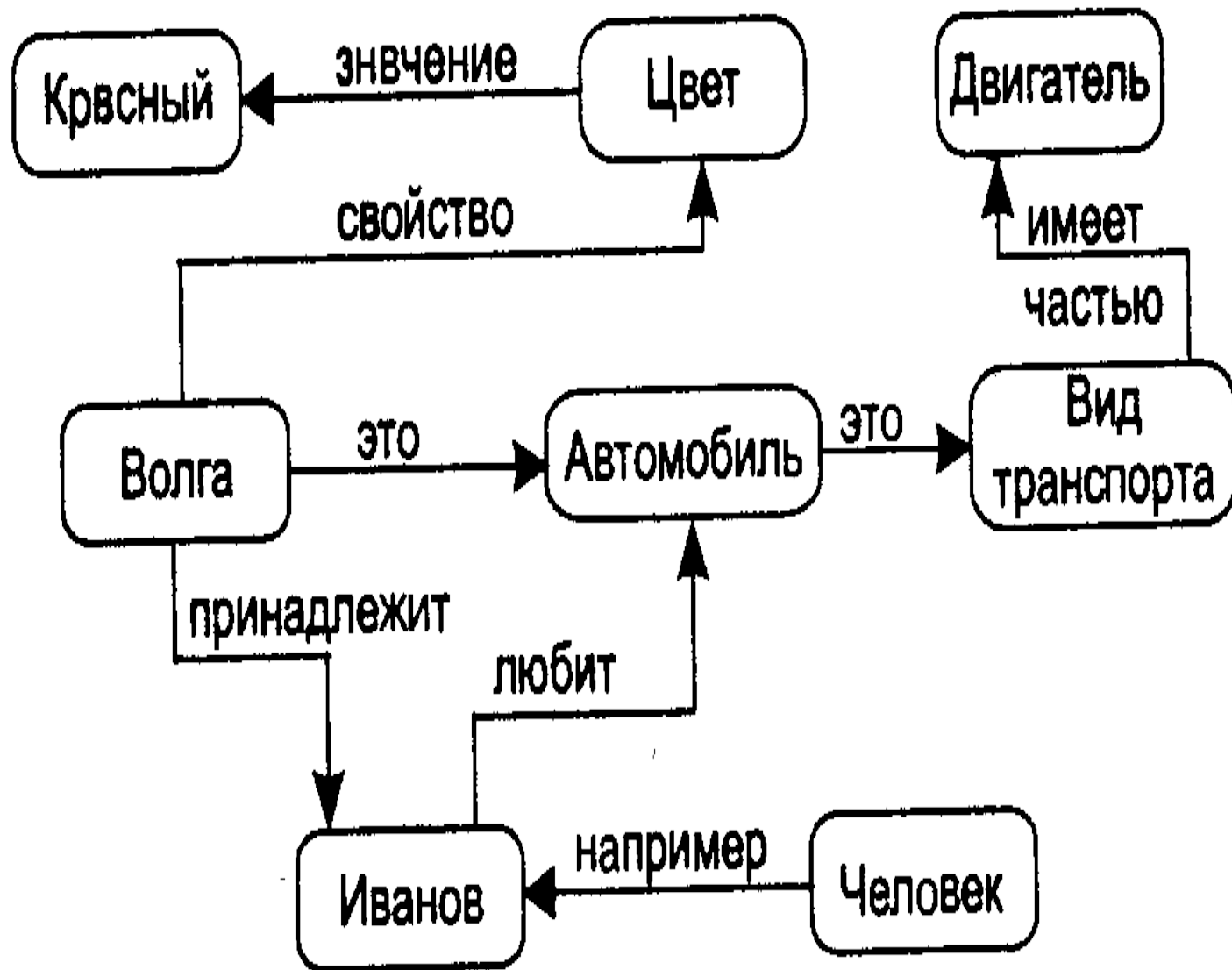
Классификация отношений в семантических сетях

- количественные (больше, меньше, равно...);
- пространственные (далеко от , близко от, за, под, над ...);

Классификация отношений в семантических сетях

- временные (раньше, позже, в течение...);
- атрибутивные связи (иметь свойство, иметь значение);
- логические связи (И, ИЛИ, НЕ);
- лингвистические связи и др.

Пример семантической сети



Достоинства семантических сетей

Семантическая сеть как модель представления знаний была предложена американским психологом Куиллианом. Основным ее преимуществом является то, что она более других соответствует современным представлениям об организации долговременной памяти человека.

Недостаток семантических сетей

Недостатком этой модели является сложность организации процедуры поиска вывода на семантической сети.

Использование семантических сетей

Для реализации семантических сетей существуют специальные сетевые языки, например NET , язык реализации систем SIMER+MIR [Осипов, 1997] и др.

Существуют экспертные системы, использующие семантические сети в качестве языка представления знаний - PROSPECTOR, CASNET, TORUS.

Фреймовая модель

Термин *фрейм* (от английского frame, что означает "каркас" или "рамка") был предложен *Марвином Минским*, одним из пионеров ИИ, в 70-е годы для обозначения структуры знаний для восприятия пространственных сцен. Эта модель, как и семантическая сеть, имеет глубокое психологическое обоснование.

Определение фрейма

Фрейм – это абстрактный образ для представления некоего стереотипа восприятия.

Фрейм – это формализованная модель восприятия образа.

Различают фреймы-образцы, или прототипы, хранящиеся в базе знаний, и фреймы-экземпляры, которые создаются для отображения реальных фактических ситуаций на основе поступающих данных.

Виды фреймов

Модель фрейма является достаточно универсальной, и позволяет отобразить все многообразие знаний о мире через:

- фреймы-структуры, использующие для обозначения объектов и понятий (заем, залог, вексель);
- фреймы-роли (менеджер, кассир, клиент);
- фреймы-сценарии (банкротство, авария, рабочий режим устройства) и др.

Структура фрейма

Структура фрейма может быть представлена как список свойств:

(ИМЯ ФРЕЙМА:

(имя 1 слота: значение 1 слота),

(имя 2 слота: значение 2 слота),

.

.

.

(имя N слота: значение N слота)).

Пример фрейма

(Список работников:

(Фамилия : значение слота 1)

(Год рождения: значение слота 2)

(Специальность: значение слота 3)

(Стаж: значение слота 4))

Структура фрейма

Ту же запись можно представить в виде таблицы, дополнив ее 2-мя столбцами:

Имя фрейма			
Имя слота	Значение слота	Способ получения значения	Присоединенная процедура

Структура фрейма

В таблице дополнительные столбцы предназначены для способа описания получения слотом значения и возможного присоединения к тому или иному слоту спец. процедур, что допускается в теории фреймов. В качестве значения слота могут выступать имя другого фрейма, так образуются сети фреймов.

Получение значения слотом фрейма

Существует несколько способов получения слотом значений во фрейме-экземпляре:

- по умолчанию от фрейма-образца (Default-значение);
- через наследование свойств от фрейма, указанного в слоте АКО;
- по формуле, указанной в слоте;
- через присоединенную процедуру;
- явно из диалога с пользователем;
- из базы данных.

Получение значения слотом фрейма

Важнейшим свойством теории фреймов является заимствование из теории семантических сетей - так называемое *наследование свойств*. И во фреймах, и в семантических сетях наследование происходит по *АКО-связям* (*A-Kind-Of = это*). Слот АКО указывает на фрейм более высокого уровня иерархии, откуда неявно наследуются, то есть переносятся, значения аналогичных слотов.

Достоинства и недостатки фреймовой модели

Основным преимуществом фреймов как модели представления знаний является то, что она отражает концептуальную основу организации памяти человека, а также ее гибкость и наглядность.

Фреймовые языки представления знаний

Специальные языки представления знаний в сетях фреймов FRL (Frame Representation Language), KRL (Knowledge Representation Language), фреймовая "оболочка" Каппа и другие программные средства позволяют эффективно строить промышленные ЭС. Известны такие фрейм-ориентированные экспертные системы, как ANALYST, МОДИС, TRISTAN, ALTERID.