

Инженерия знаний

Преподаватель:

Шошина Ксения Владимировна, к.т.н.,
доцент кафедры информационных систем и технологий,
Ауд. 10212



Цель дисциплины

Целями изучения дисциплины «Инженерия знаний» являются

- предоставление студентам систематизированных знаний о методах, языках, системах и технологиях, используемых при разработке систем, основанных на знаниях, в частности, экспертных систем;
- приобретение студентами навыков, необходимых инженеру знаний при построении систем, основанных на знаниях;
- формирование у студентов аналитических способностей, которые бы позволили им делать обоснованный выбор изученных методов и технологий при решении различного класса задач из заданной предметной области.

Краткое содержание

- **Раздел 1. Модели и средства представления знаний.**

Введение в инженерию знаний. Предмет и метод инженерии знаний. Основные модели представления знаний. Логическая модель представления знаний. Сетевая модель. Понятие семантической сети. Фреймы. Продукционная модель. Структура программной системы. Представление нечетких знаний. Нечеткий вывод. Онтологии. Визуальное представление знаний. Интеллектуальные и концептуальные карты.

- **Раздел 2. Экспертные системы.**

Общее понятие экспертных систем (ЭС). Системы объяснений в ЭС. Принципы построения. Основные достоинства и недостатки. Базы знаний экспертных систем. Представление знаний о предметной области. Экспертные игры.

7 СЕМЕСТР

1	Модели и средства представления знаний. Введение в инженерию знаний. Предмет и метод инженерии знаний.		4							10
2	Формальные модели представления знаний.		4			10				15
3	Онтологии. Основные определения. Визуальное представление знаний.		4			8				30
4	Экспертные системы. Основные особенности ЭС. Экспертные игры.		6			12				37
	Консультации	4							4	
	Форма промежуточного контроля					зачет				
	Всего	144	18			30		4		92

Давайте знакомиться!



Инженерия знаний БИ 2019



Открытая группа

1 участник

Определения

- **Инженерия знаний** представляет собой совокупность моделей, методов и технических приемов, нацеленных на создание систем, которые предназначены для решения проблем с использованием знаний.
- Фактически инженерия знаний – это теория, методология и технология, которые охватывают методы добычи, анализа, представления и обработки **знаний экспертов**.

Определения

- Представление знаний, их обработка и использование, рассматриваемые применительно к конкретной прикладной области, являются **предметом инженерии знаний.**

Процессы ИЗ

На высоком уровне процесс инженерии знаний состоит из двух:

- 1. **Извлечение знаний** – преобразование «сырых знаний» в организованные.
- 2. **Внедрение знаний** – преобразование организованных знаний в реализованные.

Понятие ИИ

Сущностью искусственного интеллекта (ИИ) можно считать научный анализ и автоматизацию интеллектуальных функций человека. Однако для большинства проблем общая реальность – трудность их **машинного** воплощения. Исследования по ИИ позволили утвердиться во мнении, что подлинно необходимым для решения проблем являются **знания экспертов**, т.е. если создать систему, способную запоминать и использовать знания экспертов, то она найдет применение в практической деятельности.

История возникновения термина

- В конце 1960-х и начале 1970-х гг. под руководством **Эдварда Фейгенбаума** в Стенфордском университете США были разработаны ЭС (DENDRAL, а позднее – MYCIN).
- Слово «engineering» в английском языке означает искусную обработку предметов, изобретение или создание чего-либо.
- Следовательно, работу по оснащению программ специальными экспертными знаниями из проблемной области, выполняемую человеком либо компьютером (программой), также можно назвать «инженерией»



Базы знаний

- **Знания** - совокупность сведений, понятий, представлений о чем-либо, полученных, приобретенных, накопленных в результате учения, опыта, в процессе жизни и т.д. и обычно реализуемых в деятельности.

Этапы обработки знаний



Базы знаний

- База знаний - это семантическая модель, описывающая предметную область и позволяющая отвечать на такие вопросы из этой предметной области, ответы на которые в явном виде не присутствуют в базе.
- База знаний является основным компонентом систем Искусственного интеллекта и Экспертных систем.

Задачи инженерии знаний

- Анализ предметной и проблемной областей.
- Приобретение знаний.
- Выявление источников знаний.
- Автоматизация процесса сбора знаний.
- Представление знаний.
- Модели представления знаний.
- Выбор способа представления знаний.
- Поиск и хранение знаний.

Анализ предметной и проблемной областей.

- **Предметная область** - сфера человеческой деятельности, выделенная и описанная согласно установленным критериям.
- В описываемое понятие должны входить сведения об ее элементах, явлениях, отношениях и процессах, отражающих различные аспекты этой деятельности.
- В описании предметной области должны присутствовать характеристики возможных воздействий окружающей среды на элементы и явления предметной области, а также обратные воздействия этих элементов и явлений на среду.

Проблемная область - комплексное понятие, включающее предметную область, решаемые задачи, цели, возможные стратегии и эвристики.

Предметную область можно определить как объект или, например, производственную систему со всем комплексом понятий и знаний о ее функционировании. При исследовании проблемной области необходимы знания о задачах, решаемых в производственной системе, и стоящих перед ней целях.

Приобретение знаний.

Классификация этапов обучения, соответствующих способностям компьютеров к формализации знаний:

- **А. Получение информации без логических выводов.**
 - 1. Ввод программ.
 - 2. Ввод фактических данных.
- **Б. Получение извне информации, уже представленной в виде знаний.**
 - 1. Получение готового набора знаний, представленных во внутреннем формате.
 - 2. Получение знаний, представленных во внутреннем формате, в режиме диалога.
 - 3. Получение знаний, представленных во внешнем формате, и их понимание.
- **В. Обучение по примерам.**
 - 1. Параметрическое обучение.
 - 2. Обучение на основе выводов по аналогии.
 - 3. Обучение на основе выводов по индукции – эвристическое обучение.
- **Г. Приобретение знаний на метауровне.**

Выявление источников знаний.

Метод	Описание
Наблюдение на рабочем месте	Наблюдать за экспертом, решающим реальные задачи на своем рабочем месте.
Обсуждение задач	Выявить виды данных, знаний и процедур, необходимых для решения конкретных задач.
Описание задач	Попросить эксперта описать прототипную задачу для каждой категории возможных ответов.
Анализ задачи	Представить эксперту ряд реалистических задач для решения вслух с целью выявить логические основания конкретных шагов рассуждения.
Доводка системы	Попросить эксперта предоставить вам несколько задач для решения и с использованием правил, выявленных во время интервью.
Оценивание системы	Попросить эксперта проверить работу системы и подвергнуть критике правила и структуру управления прототипной системой.
Проверка системы	Предоставить примеры, решенные экспертом и прототипом системы, другим независимым экспертам для сравнения и оценки.



СЕВЕРНЫЙ
ФЕДЕРАЛЬНЫЙ
УНИВЕРСИТЕТ
имени М. В. ВОЛГОДИНСКОГО

Автоматизация процесса сбора знаний.

- Автоматизация извлечения знаний и запись их в БЗ. Неавтоматизированный сбор знаний специалистов трудоемкий процесс. В развитых интеллектуальных системах предусматриваются вспомогательные средства для приобретения знаний.
- Автоматическая структуризация неформальных знаний, доступных в Интернет через распределенную гипермедиа систему – Web.

Представление знаний.

Существует также ряд общих для всех СПЗ проблем. К ним можно отнести, в частности, проблемы:

- приобретения новых знаний и их взаимодействие с уже существующими;
- организации ассоциативных связей;
- выбора диапазона в размере элементов представления, связанной с тем, насколько «детально могут быть описаны объекты и события, и какая часть внешнего мира может быть представлена в конкретной системе»;
- неоднозначности и выбора семантических примитивов;
- модульности и понимания;
- явности знаний и доступности;
- выбора соотношения декларативной и процедурной составляющих представления, что влияет на экономичность системы, полноту, легкость кодировки и понимания.

Модели представления знаний.

- **Декларативная модель** представления знаний основывается на предположении, что проблема представления некоей предметной области решается независимо от того, как эти знания потом будут использоваться.
- Поэтому модель как бы состоит из двух частей: **статических описательных структур знаний и механизма вывода**, оперирующего этими структурами и практически независимого от их содержательного наполнения.
- В декларативных моделях не содержатся в явном виде описания выполняемых процедур. Эти модели представляют собой множество утверждений. Предметная область представляется в виде синтаксического описания ее состояния.
- Вывод решений основывается в основном на процедурах поиска в пространстве состояний.

- **В процедурном представлении** знания содержатся в процедурах небольших программах, которые определяют, как выполнять специфичные действия (как поступать в специфичных ситуациях).
- При этом можно не описывать все возможные состояния среды или объекта для реализации вывода. Достаточно хранить некоторые начальные состояния и процедуры, генерирующие необходимые описания ситуаций и действий.
- При процедурном представлении знаний семантика непосредственно заложена в описание элементов базы знаний, за счет чего повышается эффективность поиска решений.

Выбор способа представления знаний.

- При использовании **логики предикатов первого порядка** (дедуктивной логики) БЗ может рассматриваться как совокупность логических формул, которые обеспечивают частичное описание проблемной среды.
- **Семантические сети** позволяют описывать свойства и отношения объектов событий, понятий, ситуаций или действий с помощью направленного графа, состоящего из вершин и помеченных ребер.
- **Фреймы** представляют собой декларативно-процедурные структуры. Во многих фреймовых структурах возможна реализация наследственных отношений, при которых объекты могут наследовать атрибуты более абстрактных объектов. Такая форма организации знаний позволяет экономить объем памяти.

- **Продукционные модели** (основанные на правилах вида Если-То) являются наиболее популярным способом представления знаний. При организации знаний с использованием продукционных моделей в БЗ содержатся правила продукций, а в БД содержится информация, которая отображает текущее состояние решаемой задачи. Инициализацию необходимого правила осуществляет блок управления.
- Большие трудности возникают при создании моделей **нечетких знаний**.
- Формализация таких знаний осуществляется на основе **теории нечетких множеств**. Развиваются также модели на основе искусственных нейронных сетей (ИНС), многоагентных систем, генетических алгоритмов и другие модели представления и обработки знаний.

Поиск и хранение знаний.

- Поиск и хранение необходимых знаний связаны с понятием **корпоративной памяти**, которая по аналогии с человеческой памятью позволяет пользоваться предыдущим опытом и избегать повторения ошибок, что является пока достаточно труднореализуемым на практике.
- **Корпоративная память** хранит неоднородную информацию из различных источников и делает ее доступной пользователям для решения корпоративных задач.
- Большую популярность в последнее время приобретают **онтологии**.

Спасибо за внимание!