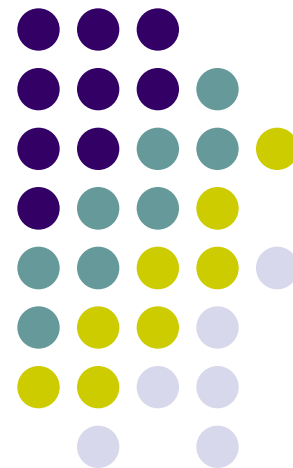


МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ И ОНТОЛОГИИ:

ВВЕДЕНИЕ

Большакова Елена Игоревна,
Груздева Надежда Валерьевна,
Баева Наталия Валерьевна



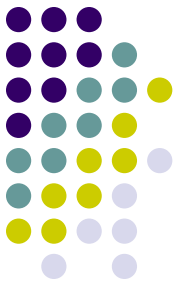
СОДЕРЖАНИЕ



1. Управление знаниями: цели, аспекты
2. Искусственный интеллект
и инженерия знаний
3. Данные *Vs* Знания : особенности
4. Разновидности знаний
5. Понятие модели и языка ПЗ,
основные модели представления знаний
6. Онтологии и их языки
7. Учебные работы в рамках курса МПЗиО
8. Литература

УПРАВЛЕНИЕ ЗНАНИЯМИ

«Знания – сила» (Ф. Бэкон)



- Задача УЗ и термин (*Knowledge Management*) возникли
 - в последние 25-30 лет (с 80-х гг.)
 - на стыке различных дисциплин (экономика, психология, менеджмент, управление персоналом, информатика)
 - в связи с накоплением используемой информации и возникновением сети Интернет
- Причины возникновения и развития концепции УЗ:
 - увеличение объема данных (в первую очередь, неструктурированных), рост сложности решаемых задач ПО
 - высокая динамика развития бизнеса, изменения рынка
 - в больших проектах требуется коллективный опыт работы, принятие решений на основе многих факторов
- Нет единого подхода к задаче, общепринятого определения термина

УПРАВЛЕНИЕ ЗНАНИЯМИ: ОПРЕДЕЛЕНИЕ, ЦЕЛИ

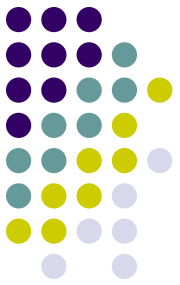


- УЗ – это процесс извлечения, распространения и эффективного использования знаний (Davenport, 1994)
- УЗ – совокупность методов, приемов, программно-технологических средств для обеспечения выявления, сбора, накопления, интеграции и хранения, распространения и применения ценной информации в ходе совместной деятельности людей

Цели:

- более эффективное функционирование коллективов
- повышение качества выпускаемых продуктов
- сохранение, тиражирование и передача опыта работы
- быстрая практическая выработка новых идей и решений
- Информация и знания, опыт и навыки сотрудников составляют основу **интеллектуального капитала** компаний, организаций, коллективов людей

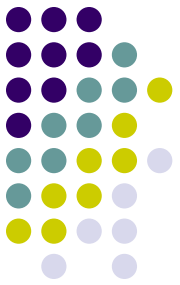
ЗНАНИЯ: ХАРАКТЕРИСТИКИ



Знания могут быть явные/неявные:

- **Явные** представлены в первую очередь информационными ресурсами:
 - документы (текстовые: электронные, письменные)
 - базы данных и знаний
 - e-мейлы
 - видео и изображения
- **Неявные** знания людей, сотрудников:
 - их интуиция
 - невербализованный экспертный опыт и навыки
 - незафиксированные организационные и управленческие знания,
 - нормы деятельности и установки, мнения
- В целом обеспечивают эффектив. принятие решений
- Важно: обмен знаниями, поддержание их целостности

УПРАВЛЕНИЕ ЗНАНИЯМИ: КОМПОНЕНТЫ И АСПЕКТЫ



- Компоненты системы УЗ
 - ❖ Люди (сотрудники компаний, организаций)
 - ❖ Организационные процессы (бизнес-процессы, совместная работа и обмен информацией)
 - ❖ Технологии (в том числе корпоративные программные системы, например, системы типа Workflow)
- Разные аспекты УЗ:
 - Информационный (информационные технологии)
 - Организационный (структуры и процессы)
 - Гуманитарный / экологический (люди и культура: взаимодействия сотрудников, ценности, установки и т.п.)
- В курсе МПЗиО – информационный аспект: явные знания, их формализация, систематизация, классификация, обработка – область **инженерии знаний** (раздел **ИИ**)

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ПАРАДИГМЫ И ЗАДАЧИ



- Термин ИИ (*Artificial Intelligence*) ввел Дж.Маккарти в 1956 г. : моделирование интеллектуальных функций
- Научное направление ИИ возникло на стыке наук: когнит. психология, математика, Computer Science, лингвистика
- Различные цели (парадигмы) ИИ →
разные принципы построения систем:
 - Попытка моделирования процесса человеческого мышления
 - Примерно с 70 гг. основная цель – функциональное моделирование для помощи человеку в решении сложных и частных задач
- **Инженерия знаний** (представление знаний), 70-е гг. –
традиционный раздел ИИ
(по сути предполагается использование знаний, как результата интеллектуальной деятельности)

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ИНЖЕНЕРИЯ ЗНАНИЙ

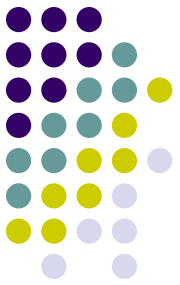


Инженерия знаний охватывает модели и методы извлечения, структурирования, формализации знаний для использования в интеллектуальных информационных системах.

Включает:

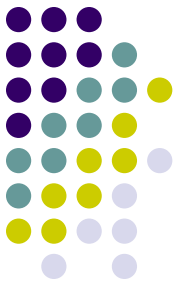
- **Извлечение/приобретение** знаний
 - первоначальное – из текстов, от человека (*eliciting*)
 - дополнительное – для расширения базы знаний системы (*acquisition*)
- **Представление знаний (ПЗ)**
 - Структурирование и формализация
 - Модели ПЗ: эвристические и формальные
- «Знания с трудом передаются от человека к человеку, еще труднее передать их машине, тем более заставить машину с толком ими воспользоваться» (Г.С. Цейтин)

ИНЖЕНЕРИЯ ЗНАНИЙ: РАЗВИТИЕ



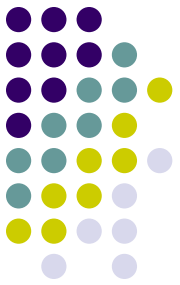
- 70-е годы – различные формальные модели представления знаний
- 80-е годы – модели предметных областей и их использование в экспертных системах
- 90-е годы – понятийные модели, **онтологии**
- 00-е годы – языки онтологий, **онтологический инжиниринг**
 - **Онтология** – концептуальная модель задачи/области, средство структурирования информационного контента
 - Онтологии – основа концепции **Semantic Web**
 - **Семантический Веб** – расширение существующего, конечная цель – позволить пользователю быстро получать ответы на сложные запросы.
 - **Semantic Web** – основная парадигма и технология Управления Знаниями будущего.

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ: ДАННЫЕ И ЗНАНИЯ



- «Знания – проверенный практикой результат познания действительности» (Философский словарь)
- Близкие термины: информация – данные – знания
В чем отличие? Что представляют собой знания?
...информация, представленная в тексте...
...данные, которые обрабатывает программа...
- Нет четкого разграничения, однако более узкое значение
- Знания – сложноорганизованные данные, для которых характерно:
 - Внутренняя интерпретируемость
 - Структурированность
 - Семантическая связность единиц
 - Активность

ЗНАНИЯ: ИНТЕРПРЕТАЦИЯ, СТРУКТУРИРОВАННОСТЬ



- Внутренняя (явная) интерпретируемость, мнемоничность
Сравнить: запись
*record { имя: Федор; фамилия: Сидоров;
возраст: 30; должность: инженер }*
и ее двоичный аналог (последовательность цифр):
010100101100011100...
- Структурированность: вложенность одних информационных единиц в другие
 - Аналог: структуры данных в программировании, например: массив записей
 - Но это только один вид связи информационных единиц, а в действительности есть много других,
семантических связей

ЗНАНИЯ: СЕМАНТИЧЕСКИЕ СВЯЗИ

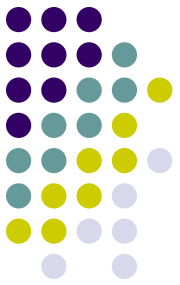


Семантические связи/отношения единиц знаний:

- Общелогические (род-вид, часть-целое)
- Причинно-следственные (каузальные)
- Пространственные (например: «находиться рядом»)
- Временные (например: «происходить одновременно»)
- Узко-предметные (например: родственные связи)
- Функциональные (например: год рождения-возраст)
- *Ассоциативные связи* (например: жених-свадьба)

Когн. психология: объективно обоснованная гипотеза: когнитивные структуры человека погружены в некое пространство, метрика которого характеризует семантическую близость понятий, фактов, явлений.

ЗНАНИЯ: АКТИВНОСТЬ



- Активность – роль при решении задач
Традиционно: данные пассивны,
команды активны и инициируют процессы обработки
- Как показывают исследования когнитивных структур человека, ситуация как бы противоположная:
имеющиеся/полученные знания, их состав и структура способствуют актуализации процессов мышления.
Пример: потеря/кража мобильного телефона...
- Компьютерные программы: представление **процедурных знаний**. В ходе развития программирования:
 - накопление знаний о способах решения задач
 - движение от представления данных к знаниям, от процедурного программирования к объектно-ориентированному.

ЗНАНИЯ: ЕЩЕ ОДИН АСПЕКТ



- Знания – сложноорганизованные данные
о существенном фрагменте действительности
- **Предметная область** – фрагмент действительности, характеризующийся совокупностью объектов (сущностей), их свойств и отношений.
- **Проблемная область (ПО)** –
предметная область + решаемые в ней задачи
(в одной предметной области может быть несколько ПО)
 - Родственные отношения + установление родства
 - Флора Земли + распознавание / классификация растения
 - Детские болезни + диагностика / фармацевтика
- **База знаний** – модуль системы ИИ, в котором знания о ПО представлены в явной форме и отделены от прочих знаний системы, причем форма их представления обычно допускает разнообразное, многократное применение.

Отличие от **Базы Данных**?

ВИДЫ ЗНАНИЙ



Разнородность знаний → их характеристика
по нескольким видам/ аспектам:

- Фактографические Vs. Понятийные
 - Декларативные Vs. Процедурные
 - Глубокие Vs. Поверхностные
 - Общие Vs. Специальные (проблемно-ориентированные)
 - Предметные Vs. **Метазнания**: знания о предметных знаниях и умениях для их эффективного применения
 - Эмпирические Vs. Рациональные (по способу познания)
 - Общеизвестные Vs. Индивидуальные (по носителю)
 - Объективные Vs. Субъективные (по достоверности, истинности)
- ❖ Разделение (по каждому аспекту) – не абсолютно, и до определенной степени условно (т.н. **биполярные шкалы**)

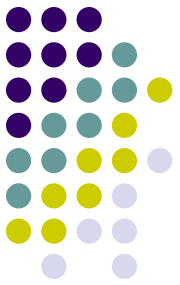
ЗНАНИЯ: ПОНЯТИЙНЫЕ И ФАКТОГРАФИЧЕСКИЕ



Аспект фактуальности (конкретности)

- **Фактографические знания** (конкретные сведения) – количественные и качественные характеристики различных объектов, явлений и их элементов.
- **Понятийные (концептуальные) знания** – набор понятий некоторой ПО, а также их свойства и взаимосвязи.
- Две части **БЗ**: – фактографическая (*экстенсиональная*)
– концептуальная, понятийная (*интенсиональная*)
- Относительность разделения: в зависимости от ПО один и тот же факт, может быть отнесен к разным частям.
- Примеры: *А.С. Пушкин родился в 1799 г.*
Кресло включает подлокотники. Луна – спутник Земли.
Тигр – хищник, и питается мясом.
Желток яйца – желтого цвета.

ЗНАНИЯ: ДЕКЛАРАТИВНЫЕ И ПРОЦЕДУРНЫЕ



Аспект активности/пассивности

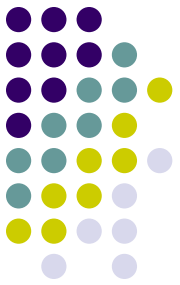
- **Процедурные знания** (как?) – знания о стратегиях, методах, алгоритмах, эвристиках решения задач.
- **Декларативные знания** (что?) – набор сведений (фактов, понятий), относящихся к некоторой ситуации/ПО.
- Декларативные знания (описания) не несут в себе информации о том, как они будут использованы, поэтому имеют более общий характер (годны для разных целей).
- Примеры описаний: «как доехать» *Vs.* «где расположен»
- Деление на процедурные и декларативные знания опять же условно, хотя различия прослеживаются:
 - в ЕЯ – изъявительное/повелительное наклонение
 - в программировании – данные и алгоритмы
но, напр.: язык Пролог *`mother(X):- parent(X,_),woman(X).`*

ЗНАНИЯ: ГЛУБОКИЕ И ПОВЕРХНОСТНЫЕ



- **Глубокие знания** – основные закономерности, аксиомы, понятия, факты о рассматриваемой ПО, общие принципы развития ситуаций и процессов ПО и свойства этих процессов.
- Глубокие знания основаны на некоторой строгой общепринятой теории.
- **Поверхностные знания** – эвристики и некоторые закономерности, устанавливаемые опытным путем, используемые обычно при отсутствии общих теорий и методов.
- Поверхностные знания основаны на проведенных опытах и наблюдениях.

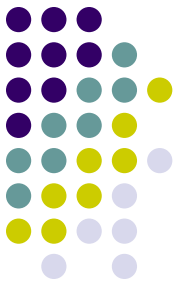
ВИДЫ ЗНАНИЙ И ФОРМЫ ПРЕДСТАВЛЕНИЯ



Разные виды знаний стимулировали разные формы и методы своего представления

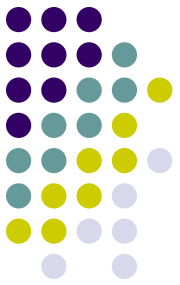
- фактографические – в форме Баз Данных
- процедурные – в форме
 - языков программирования,
 - программных библиотек,
 - прикладных программ
- понятийные – специализированные языки логики и ИИ

ПРЕДСТАВЛЕНИЕ И ОБРАБОТКА ЗНАНИЙ



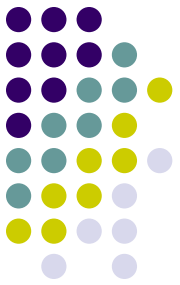
- Некоторая совокупность знаний ценна не сама по себе, а возможностью её применения → вместе с уточнением способа ПЗ надо обсуждать и возможные для этого способа процедуры оперирования знаниями.
- Обработка/оперирование знаниями:
 - поиск нужной информации, предполагающий в том числе **вывод новых знаний** из уже имеющихся
 - преобразование: абстрагирование (свертка)
 - проверка полноты и непротиворечивости знаний
- Необходимость вывода знаний:
хранить в системе все факты не только неэффективно, но и невозможно (число возможных взаимоотношений объектов даже в простой ПО бесконечно)
Вывод – за счет интенсификации (понятийной) части БЗ

МОДЕЛЬ И ЯЗЫК ПЗ



- **Метод (модель) представления знаний** – совокупность взаимосвязанных средств формального описания знаний и оперирования ими
 - Модель: структурное или функциональное подобие
 - ИИ: эвристическое моделирование
- **Язык представления знаний** – формальный искусственный язык для описания знаний и операций с ними в рамках выбранной модели.
- В рамках одной модели может быть несколько различных языков представления знаний.
- Представление знаний – это выражение на ЯПЗ свойств различных объектов и явлений ПО, обычно реализует это *инженер по знаниям*

ОСНОВНЫЕ МОДЕЛИ ПЗ



(порядок исторический)

- Логическая модель
 - единица знаний – логическая формула
- Сетевая модель
 - семантическая сеть (нагруженный граф)
- Фреймовая модель
 - единица знания – *фрейм*
- Продукционная модель
 - единица знаний – *правило продукции*

Сравнение основных моделей ПЗ:

- ❖ по выразительным возможностям (их ограничения)
- ❖ по возможности/удобству представления различных видов знаний

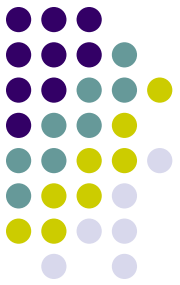
ЛОГИЧЕСКАЯ МОДЕЛЬ ПЗ



- Основная идея: использование логического языка, например, языка предикатов
(язык предикатов возник из ЕЯ)
- Исчисление предикатов первого порядка
$$\forall x (\text{Студент}(x) \wedge \text{Старший}(\text{курс}(x)) \rightarrow \exists y \text{ Научный_руководитель}(y, x))$$

(формула ИП может быть записана на языке Пролог)
- Особенность: **механизм** (процедура) **логического вывода** новой информации (знаний)
- **Дескриптивные логики**
 - ограничение выразительной мощности языка предикатов для достижения эффективности вывода
 - дополнительные синтаксические средства

СЕТЕВАЯ И ФРЕЙМОВАЯ МОДЕЛИ ПЗ



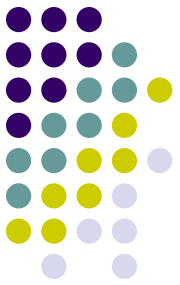
Сетевая модель

- Основная идея – рассматривать ПО как набор объектов (сущностей) и связей между ними
- Связи вершин: семантические (общелогические и др.)
- **Семантическая сеть** – ориентированный граф из поименованных вершин и ребер
- Единица знаний – ребро и 2 связываемые им вершины

Фреймовая модель – развитие, разновидность сетевой

- Единица знания – *фрейм*, упрощенно: структура, объединяющая основные свойства/атрибуты некоторой сущности (фрейм – набор атрибутов)

ПРОДУКЦИОННАЯ МОДЕЛЬ

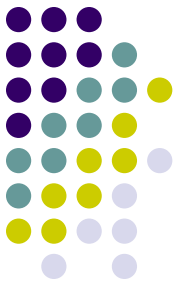


Пример правила продукции на языке *CLIPS*
(правило из БЗ экспертной системы):

```
(defrule anemia          ;распознавание болезни
(declare (salience 115)) ;приоритет правила
(not(suggest ?))
(name headache)          ;левая часть правила - симптомы
(name weaknesses)
(name paleness)
(name faint)
=>                        ; правая часть: предположение о диагнозе
(assert(suggest "This is anemia, because you
  have headache,
weaknesses, paleness and faint"))
)
```

CLIPS: Парадигма ситуационного программирования

ОНТОЛОГИИ И ИХ ЯЗЫКИ



- Онтология – инвентарь сущностей (объектов) некоторой ПО, их описаний и описаний их связей
- Кроме описания понятий и связей онтология может включать правила их использования, ограничивающие их значения в рамках конкретной ПО
- Типы онтологий: формальные, лингвистические
- Специальные формальные языки онтологий:
 - Дескриптивные логики (семейство логик со своими выразительными средствами)
 - Языки *Semantic Web: RDF, RDFS, OWL*
- Программные средства разработки онтологий
 - Редакторы онтологий
 - Конверторы для разных языков описания онтологий



СПАСИБО ЗА ВНИМАНИЕ!

УЧЕБНЫЕ РАБОТЫ КУРСА



- Лекции (в форме презентаций)
- Самостоятельные аудиторные работы (на 5-15 минут)
- Домашние работы (без программирования) – на формализацию и представление знаний (на 1-2 недели)
- Обязательные 3 практические домашние задания, с использованием специализированных компьютерных систем (включают формализацию и программирование)
- ❖ Накопительная балльная система определения **оценки за курс** (получение ее без экзамена, в мае)
- ❖ Контрольная /экзамен в июне

ЛИТЕРАТУРА



- Добров Б.В. и др. Онтологии и тезаурусы: модели, инструменты, приложения – М.: Интернет-университет информационных технологий, 2009, 176 с.
- Джарратано Дж., Райли Г. Экспертные системы. Принципы разработки и программирование – М.: ООО «И.Д. Вильямс», 2007.
- Лапшин В.А. Онтологии в компьютерных системах. – М.: Научный Мир, 2010.
- Люгер Дж. Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. – М.: Издательский дом «Вильямс», 2005.
- Новиков Ф.А. Дискретная математика для программистов: Учебник для вузов. 3-е изд. – СПб: Питер, 2009.
- Рассел С., Норвиг П. Искусственный интеллект: современный подход, 2-е изд. // Пер. с англ. – М.: Издательский дом «Вильямс», 2006.
- Частиков А.П., Гаврилова Т.А., Белов Д.Л. Разработка экспертных систем. Среда CLIPS. – СПб.: БХВ-Петербург, 2003.