

Лекция №8

ТЕХНОЛОГИЯ БАЗ ЗНАНИЙ (БЗ) И СИСТЕМЫ УПРАВЛЕНИЯ ЗНАНИЯМИ (СУЗ)

Создание БЗ и в теории, и в практике ИИ сегодня является проблемой такой же важности, как в свое время в информационных технологиях проблема создания БД.

Под БЗ понимается семантическая модель, предназначенная для представления в ЭВМ знаний, накопленных человеком в определенной предметной области.

На технологическом уровне БЗ рассматривается как хранилище (репозиторий) сложно структурированных информационных единиц (знаний).

БЗ подразделяются на **замкнутые** и **открытые**.

Интерпретация содержимого **замкнутой БЗ** в процессе функционирования включающей ее интеллектуальной системы не изменяется. Логический вывод в такой БЗ эквивалентен выводу в формальной системе и обладает свойством монотонности.

Противоположные черты присущи **открытой БЗ**. Охватывающая ее интеллектуальная система может пополнять и модифицировать содержимое БЗ, а также удалять знания из нее. Вывод в открытой БЗ является немонотонным.

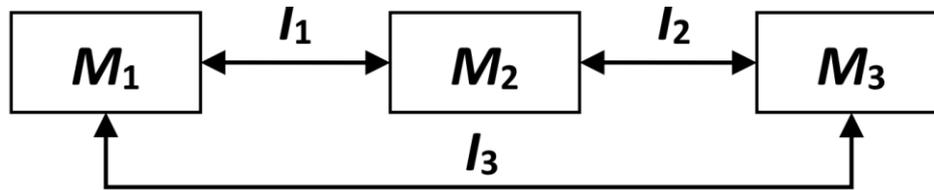
БЗ



Говоря о БЗ, всегда соотносят ее со знаниями о некоторой предметной области. При этом под предметной областью может пониматься и некоторый класс решаемых задач.

По аналогии с технологией БД будем различать собственно **информационное хранилище знаний (БЗ)** и **систему управления БЗ (СУБЗ)**, обеспечивающую набор типовых функций хранения и манипулирования знаниями.

Обобщенная структура БЗ



Математически БЗ представляется шестеркой $(M_1, M_2, M_3, I_1, I_2, I_3)$, где M_1 – база глубинных знаний, представляющая понятийные структуры предметной области;

M_2 – база фактов;

M_3 – база метазнаний;

I_1 – интерфейсы между M_1 и M_2 ;

I_2 – интерфейсы между M_2 и M_3 ;

I_3 – интерфейсы между M_1 и M_3 .

Обобщенная структура БЗ



База глубинных знаний M_1 состоит из двух компонентов $M_1 = (M_{11}, M_{12})$, где M_{11} — часть хранилища знаний, содержащая описания единиц знаний, образующих понятийные структуры предметной области; M_{12} — сеть фреймов над понятийными структурами.

База фактов M_2 соответствует части хранилища знаний, содержащей эмпирические данные о предметной области, параметры наблюдаемых ситуаций и т.д.

База метазнаний включает три компонента $M_3 = (M_{31}, M_{32}, M_{33})$, M_{31} — база правил для данной предметной области; M_{32} — база метаправил, метаметаправил и т.д.; M_{33} — стратегия управления правилами и метаправилами.

Обобщенная структура БЗ



Интерфейсы I_1 , I_2 и I_3 представлены парами компонентов, соответствующими направленности связей между взаимодействующими блоками БЗ:

$$I_1 = (I_{11}, I_{12}), I_2 = (I_{21}, I_{22}), I_3 = (I_{31}, I_{32}),$$

где I_{11} — интерфейс, связывающий M_1 и M_2 ;

I_{12} — интерфейс, связывающий M_2 и M_1 ;

I_{21} — интерфейс, связывающий M_2 и M_3 ;

I_{21} — интерфейс, связывающий M_3 и M_2 ;

I_{31} — интерфейс, связывающий M_1 и M_3 ;

I_{32} — интерфейс, связывающий M_3 и M_1 .

Обобщенная структура БЗ



Наиболее сложной проблемой является представление глубинных знаний (M_1).

Технология построения M_1 непосредственно связана с выбором модели представления знаний о предметной области. В настоящее время для организации M_1 используется технология объектно-ориентированных БД.

База фактов M_2 , как правило, реализуется на основе технологии реляционных БД.

Для построения базы метазнаний M_3 в последние годы все чаще используются семантические сети и онтологии.

Управление знаниями представляет собой интегрирующую интеллектуальную информационную технологию, которая объединяет в единый комплекс множество технологий, поддерживающих процессы формирования, накопления, хранения, распространения, обработки и использования знаний и данных.

Понятие «**управление знаниями**» появилось в середине 90-х годов прошлого века.

Возникновение этого направления интеллектуальных информационных технологий вызвано потребностями пользователей корпоративных ИС.

Традиционные корпоративные ИС оперируют не знаниями, а данными — документами, записями в БД, выборками, отчетами и т.п.

Управление знаниями рассматривается как совокупность процессов, управляющих созданием, распространением, обработкой и использованием знаний в рамках организации.

СУЗ должна обеспечивать:

- отражение изменений данных в корпоративной БД, характеризующих историю деятельности компании;
- извлечение, интеграцию и представление в явном виде знаний специалистов компании;
- представление информации, содержащейся в корпоративных БД, на семантическом уровне;
- анализ и извлечение знаний из данных в корпоративных БД;
- поиск и доступ к информации по смыслу;
- поддержку совместной работы с ИР специалистов компании;
- поддержку процессов формирования новых знаний.

Корпоративные знания



Корпоративные знания разделяют на три слоя:

- 1) формализованные знания, представленные в БЗ;
- 2) знания, содержащиеся в документах и БД;
- 3) профессиональные знания специалистов компании, не зафиксированные на материальных носителях.

В число задач СУЗ входит поддержка процессов:

- явного выражения (фиксации) знаний специалистов;
- формализации и автоматизированного извлечения знаний из ИР.

ИТ, которые объединяются в единый комплекс технологией СУЗ



- БД, хранилищ данных и БЗ;
- управления документооборотом;
- поддержки совместной работы с ИР;
- автоматизированного извлечения знаний из текста;
- поиска в текстовой и структурированной информации (в том числе поиска по метаданным);
- автоматической классификации и кластеризации документов;
- приобретения знаний от экспертов;
- машинного перевода;
- автоматического реферирования и аннотирования;
- интеллектуального анализа данных;
- автоматического распознавания образов;
- поддержки принятия решений;
- поддержки инновационной деятельности (формирования новых знаний).

Структура корпоративной памяти



Фундаментом СУЗ служат технологии хранилищ данных и БЗ на основе онтологического подхода.

На базе технологии хранилищ данных была сформирована **концепция корпоративной памяти (corporate memory)**, структура которой приведена в таблице.

Уровень представления информации	Вид информации		
	Документы	Данные	Знания
Онтологический	Структуры архивов	Структуры данных	Базовые онтологии
Содержательный	Отчеты, методики, инструкции	Справочники, каталоги	Правила вывода, факты
Программно реализованный	Документы (тексты, рисунки, схемы)	БД, файлы	БЗ

Внедрение СУЗ



Внедрение СУЗ в организациях, значительное число сотрудников которых занято обработкой информации, приносит ощутимый экономический эффект.

СУЗ позволяет ежедневно экономить в среднем 40-50 мин. рабочего времени одного сотрудника, что эквивалентно повышению производительности труда на 8-10%.

Общий выигрыш от использования СУЗ составляет 8-10% от соответствующего фонда заработной платы.

СУЗ, стоимость которой равна месячному фонду заработной платы, окупится примерно за год.