**1 История развития искусственного интеллекта (время и место зарождения понятия искусственный интеллект, первые направления развития искусственного интеллекта)**

Под искусственным интеллектом понимается одно из направлений информационных технологий, которое занимается изучением и разработкой систем (машин), наделенных возможностями человеческого интеллекта: способность к обучению, логическому рассуждению и так далее.

Первые работы в области ИИ начал вести в середине прошлого века Алан Тьюринг, хотя определенные идеи начали высказывать философы и математики в Средние века. В частности, еще в начале 20-го века была представлена механическое устройство, способное решать шахматные задачи.

Но по-настоящему это направление сформировалось к середине прошлого столетия. К тому времени появились первые компьютеры и алгоритмы. То есть, был создан фундамент, на котором зародилось новое направление исследований.

В 1950 году Алан Тьюринг опубликовал статью, в которой задавался вопросами о возможностях будущих машин, а также о том, способны ли они обойти человека в плане разумности.

В 1956 году группа ученых, включая Тьюринга, собралась в американском университете Дартмунда, чтобы обсудить вопросы, связанные с ИИ. После той встречи началось активное развитие машин с возможностями искусственного интеллекта и зародился термин «искусственный интеллект».

Особую роль в создании новых технологий в области ИИ сыграли военные ведомства, которые активно финансировали это направление исследований. Впоследствии работы в области искусственного интеллекта начали привлекать крупные компании.

В настоящее время в исследованиях по ИИ выделяют семь основных направлений

1. Представление знаний. Решаются задачи, связанные с формализацией и представлением знаний в памяти интеллектуальной системы (ИС).

2. Манипулирование знаниями. Строятся способы пополнения знаний на основе их неполных описаний, изучаются системы, классификации хранящихся в ИС знаний, разрабатываются процедуры обобщения знаний и формирования на их основе абстрактных понятий, создаются методы достоверного и правдоподобного вывода на основе имеющихся знаний, предлагаются модели рассуждений.

3. Общение. В круг задач этого направления входят: проблема понимания связных текстов, понимания речи и синтез речи, теория моделей коммуникации между человеком и ИС.

4. Восприятие. Включает проблемы анализ трехмерных сцен, разработку методов представления информации о зрительных образах в базе знаний, создание методов перехода от зрительных сцен к их текстовому описанию и методов обратного перехода, разработку процедур когнитивной графики, создание средств для порождения сцен на основе внутренних представлений в ИС.

5. Обучение. Предполагается, что ИС подобно человеку будут способны к обучению - решению задач, с которыми они ранее не встречались.

6. Поведение. Разрабатываются специальные поведенческие процедуры, которые позволяют ИС адекватно взаимодействовать с окружающей средой, другими ИС и людьми.

7. Интеллектуальное программирование(самое молодое). Включает в себя: языки для интеллектуального программирования (логического программирования, объектно-ориентированные языки, языки представления знаний и семантической разметки); автоматический синтез программ (дедуктивные и индуктивные методы); инструментальные средства; интеллектуальные интерфейсы; мультиагентные технологии.

**2 Интеллектуальные информационные системы (характеристика, классификация)**

Интеллектуальная информационная система – это взаимосвязанная совокупность средств, методов и персонала, имеющая возможность хранения, обработки и выдачи информации, а также самостоятельной настройки своих параметров в зависимости от состояния внешней среды (исходных данных) и специфики решаемой задачи.

Интеллектуальные информационные системы являются естественный результатом развития обычных информационных систем, сосредоточили в себе наиболее наукоемкие технологии с высоким уровнем автоматизации не только процессов подготовки информации для принятия решений, но и самих процессов выработки вариантов решений, опирающихся на полученные информационной системой данные.

ИИС особенно эффективны в применении к слабо структурированным задачам, в которых пока отсутствует строгая формализация, где при принятии решений учитываются наряду с экономическими показателями слабо формализуемые факторы — экономические, политические, социальные.

Интеллектуальные системы имеют следующие характерные признаки:

* развитые коммуникативные способности: возможность обработки произвольных запросов в диалоге на языке максимально приближенном к естественному (система естественно-языкового интерфейса — СЕЯИ);
* направленность на решение слабоструктурированных, плохо формализуемых задач (реализация мягких моделей);
* способность работать с неопределенными и динамичными данными;
* способность к развитию системы и извлечению знаний из накопленного опыта конкретных ситуаций;
* возможность получения и использования информации, которая явно не хранится, а выводится из имеющихся в базе данных;
* система имеет не только модель предметной области, но и модель самой себя, что позволяет ей определять границы своей компетентности;
* способность к аддуктивным выводам, т.е. к выводам по аналогии;
* способность объяснять свои действия, неудачи пользователя, предупреждать пользователя о некоторых ситуациях, приводящих к нарушению целостности данных.

Классификация по коммуникативным способностям:

* Интеллектуальные базы данных (можно делать выборку данных, которая может явно не храниться, а выводиться из имеющейся БД);
* Системы естественно-языкового интерфейса (СЕЯИ) (трансляцию естественно-языковых конструкций на машинный уровень представления знаний);
* Гипертекстовые системы (предназначены для поиска текстовой информации по ключевым словам в базах);
* Контекстные системы (частный случай гипертекстовых и естественно-языковых систем);
* Системы когнитивной графики (позволяют осуществлять взаимодействие пользователя ИИС с помощью графических образов).

По типу решаемых задач.

* Экспертные системы;
* Классифицирующие системы;
* Доопределяющие системы; Т
* рансформирующие системы;
* Многоагентные системы.

По способности к самообучению.

* Индуктивные системы;
* Нейронные сети;
* Системы, основанные на прецедентах;
* Информационные хранилища

**3 Экспертные системы. Основные свойства. Структура экспертной системы. Отличие статической от динамической системы**

Экспертная система – это программное средство, использующее экспертные знания для обеспечения высокоэффективного решения неформализованных задач в узкой предметной области. Основу ЭС составляет база знаний (БЗ) о предметной области, которая накапливается в процессе построения и эксплуатации ЭС. Накопление и организация знаний - важнейшее свойство всех ЭС. ЭС используются для решения так называемых неформализованных задач, общим для которых является то, что:

* задачи не могут быть заданы в числовой форме;
* цели нельзя выразить в терминах точно определенной целевой функции;
* не существует алгоритмического решения задачи;
* если алгоритмическое решение есть, то его нельзя использовать из-за
* ограниченности ресурсов (время, память).

Знания являются явными и доступными, что отличает ЭС от традиционных программ, и определяет их основные свойства, такие, как:

1) Применение для решения проблем высококачественного опыта;

2) Наличие прогностических возможностей;

3) Обеспечение институциональной памятью;

4) Возможность использования ЭС для обучения и тренировки руководящих работников

Экспертные системы в общем случае подразделяются на статические и динамические.

Стандартная статическая экспертная система состоит из следующих основных компонентов:

1) рабочей памяти, называемой также базой данных – предназначена для получения и хранения исходных и промежуточных данных решаемой в текущий момент задачи

2) базы знаний – предназначена для хранения долгосрочных данных, описывающих конкретную предметную область, и правил, описывающих рациональное преобразование данных этой области решаемой задачи;

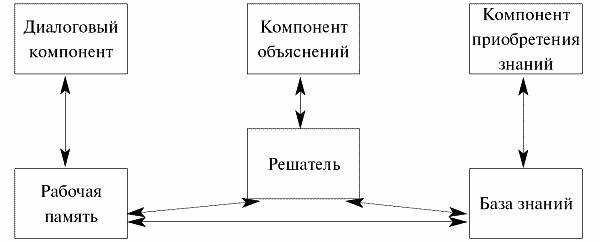
3) решателя, называемого также интерпретатором – используя исходные данные из рабочей памяти и долгосрочные данные из базы знаний, он формирует правила, применение которых к исходным данным приводит к решению задачи;

4) компонентов приобретения знаний – автоматизирует процесс заполнения экспертной системы знаниями эксперта;

5) объяснительного компонента – разъясняет, как система получила решение данной задачи, или почему она это решение не получила и какие знания она при этом использовала;

6) диалогового компонента – служит для обеспечения дружественного интерфейса пользователя как в ходе решения задачи, так и в процессе приобретения знаний и объявления результатов работы.

Струтура:



Статические экспертные системы чаще всего используются в технических приложениях, где можно не учитывать изменения окружающего среды, происходящие во время решения задачи.

В отличие от статической экспертной системы в структуру динамической экспертной системы дополнительно вводятся два следующих компонента:

1) подсистема моделирования внешнего мира;

2) подсистема связей с внешним окружением – осуществляет связи с внешним миром посредством системы специальных датчиков и контроллеров.

Помимо этого, некоторые традиционные компоненты статической экспертной системы подвергаются существенным изменениям, для того чтобы отобразить временную логику событий, происходящих в данный момент в окружающей среде.

Это главное различие между статической и динамической экспертными системами.

Пример динамической экспертной системы – управление производством различных медикаментов в фармацевтической промышленности.

**4 Целесообразность использования экспертных систем для решения поставленной задачи (условия)**

В отличие от машинных программ, используемых процедурный анализ, ЭС решает задачи в узкой предметной области на основе дедуктивных рассуждений. Такие системы часто оказываются способными найти решение задач, которые не структурированы, и плохо определены. Главным достоинством ЭС является возможность накопления, обновления знаний и сохранения их в течение длительного времени. При этом ЭС, обеспечивают независимость оценки от мнения специалистов, и позволяют повышать квалификацию специалистов компании, используя наилучшие проверенные решения.

Критерии целесообразности применения ЭС:

* данные знания должны быть надежными и мало меняться с течением времени;
* решаемые задачи должны быть узкоспециализированными, а пространство возможных решений должно быть относительно невелико;
* задачи не должны в значительной мере зависеть от общечеловеческих знаний или соображений здравого смысла;
* должен быть по крайне мере 1 эксперт способный явно сформулировать свои знания и объяснить свои методы применения этих знаний для решения задач;
* в процессе решения задачи должны использоваться формальные рассуждения.

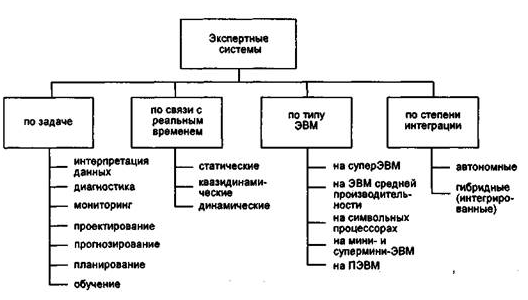
Случаи, в которых не рекомендуется использование ЭС:

* для математических задач, которые решаются путем процедурного анализа;
* для задач методы решения, которых отсутствуют.

Условие, свидетельствующие о необходимости разработки ЭС:

* нехватка специалистов, имеющих возможность затратить значительное время для оказания помощи другим специалистам;
* необходимость привлечения количества специалистов для решения небольшой задачи, поскольку ни 1 из них не обладает достаточными знаниями для решения ее самостоятельно;
* слишком большое расхождение между решениями самых квалифицированных и самых неквалифицированных специалистов;
* наличие конкурентов имеющих преимущество в силу того, что они лучше справляются с поставленной задачей.

**5 Классификация экспертных систем (по решаемой задача, по типу ЭВМ, по связи с реальным временем, по степени интеграции с другими программами)**



Классификация по решаемой задаче

Интерпретация данных. Это одна из традиционных задач для экспертных систем. Под ин­терпретацией понимается определение смысла данных, результаты которого должны быть согласованными и корректными. Обычно предусматривается многовариантный анализ данных.

Диагностика. Под диагностикой понимается обнаружение неисправности в некото­рой системе. Важной спецификой является необходимость понимания функциональной структуры ("анатомии") диагностирующей системы.

Мониторинг. Основная задача мониторинга — непрерывная интерпретация данных в реальном масштабе времени и сигнализация о выходе тех или иных параметров за допусти­мые пределы. Главные проблемы — "пропуск" тревожной ситуации и инверсная задача "ложного" срабатывания. Сложность этих проблем в размытости симптомов тревожных си­туаций и необходимость учета временного контекста.

Проектирование. Проектирование состоит в подготовке спецификаций на создание "объектов" с заранее определенными свойствами. Под спецификацией понимается весь набор необходимых документов — чертеж, пояснительная записка и т.д. Основные пробле­мы здесь — получение четкого структурного описания знаний об объекте и проблема "следа". Дня организации эффективного проектирования и, в еще большей степени, пере­проектирования необходимо формировать не только сами проектные решения, но и мотивы их принятия. Таким образом, в задачах проектирования тесно связываются два основных процесса, выполняемых в рамках соответствующей ЭС: процесс вывода решения и процесс объяснения.

Прогнозирование. Прогнозирующие системы логически выводят вероятные следст­вия из заданных ситуаций. В прогнозирующей системе обычно используется параметричес­кая динамическая модель, в которой значения параметров "подгоняются" под заданную ситуацию. Выводимые из этой модели следствия составляют основу для прогнозов с веро­ятностными оценками.

Планирование. Под планированием понимается нахождение планов действий, отно­сящихся к объектам, способным выполнять некоторые функции. В таких ЭС используются модели поведения реальных объектов с тем, чтобы логически вывести последствия плани­руемой деятельности.

Обучение. Системы обучения диагностируют ошибки при изучении какой-либо дис­циплины с помощью ЭВМ и подсказывают правильные решения. Они аккумулируют зна­ния о гипотетическом "ученике" и его характерных ошибках, затем в работе способны диагностировать слабости в знаниях обучаемых и находить соответствующие средства для их ликвидации. Кроме того, они планируют акт общения с учеником в зависимости от успе­хов ученика с целью передачи знаний.

Классификация по связи с реальным временем

Статические ЭС разрабатываются в предметных областях, в которых база знаний и интерпретируемые данные не меняются во времени. Они стабильны.

Квазидинамические ЭС интерпретируют ситуацию, которая меняется с неко­торым фиксированным интервалом времени.

Динамические ЭС работают в сопряжении с датчиками объектов в режиме ре­ального времени с непрерывной интерпретацией поступаемых данных.

Классификация по типу ЭВМ

ЭС для уникальных стратегически важных задач на суперЭВМ (Эльбрус, CRAY, CONVEX и др.).

ЭС на ЭВМ средней производительности (типа ЕС ЭВМ, mainframe).

ЭС на символьных процессорах и рабочих станциях (SUN, APOLLO).

ЭС на мини- и супермини-ЭВМ (VAX, micro-VAX и др.).

ЭС на персональных компьютерах (IBM PC, MAC II и подобные).

Классификация по степени интеграции с другими программами

Автономные ЭС работают непосредственно в режиме консультаций с пользователем для специфически "экспертных" задач, для решения которых не требуется привлекать тра­диционные методы обработки данных (расчеты, моделирование и т.д.).

Гибридные ЭС представляют программный комплекс, агрегирующий стандарт­ные пакеты прикладных программ (например, математическую статистику, линейное про­граммирование или системы управления базами данных) и средства манипулирования знаниями. Это может быть интеллектуальная надстройка над ППП или интегрированная среда для решения сложной задачи с элементами экспертных знаний.

**6 Коллектив разработчиков экспертной системы (состав, взаимодействие, зона ответственности)**

Под коллективом разработчиков (КР) будем понимать группу специалистов, ответственных за создание ЭС. Минимальный состав КР включает четыре человека (эксперт; инженер по знаниям; программист; пользователь); реально же он разрастается до 8-10 человек.  Возглавляет коллектив инженер по знаниям, это ключевая фигура при разработке систем, основанных на знаниях.

В большентве случаев к числу основных участников следует отнести саму экспертную систему, экспертов, инженеров знаний, разработчиков, средства построения ЭС и пользователей. Их основные роли и взаимоотношение приведены на рис



Экспертная система — это программное средство, использующее знания экспертов, для высокоэффективного решения задач в интересующей пользователя предметной области. Она называется системой, а не просто программой, так как содержит базу знаний, решатель проблемы и компоненту поддержки. Последняя из них помогает пользователю взаимодействовать с основной программой.

Программист разрабатывает ИС (если ИС разрабатывается заново), содержащее в пределе все основные компоненты ЭС, и осуществляет его сопряжение с той средой, в которой оно будет использовано.

Эксперт — это человек, способный ясно выражать свои мысли и пользующийся репутацией специалиста, умеющего находить правильные решения проблем в конкретной предметной области. Эксперт использует свои приёмы и ухищрения, чтобы сделать поиск решения более эффективным, и ЭС моделирует все его стратегии.

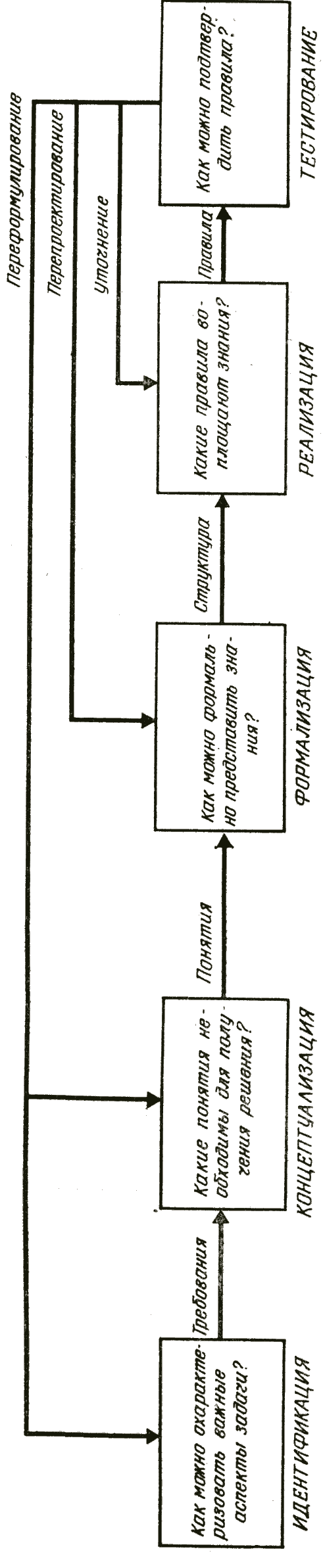
Инженер знаний — человек, как правило, имеющий познания в информатике и искусственном интеллекте и знающий, как надо строить ЭС. Инженер знаний опрашивает экспертов, организует знания, решает, каким образом они должны быть представлены в ЭС, и может помочь программисту в написании программ.

Средство построения ЭС — это программное средство, используемое инженером знаний или программистом для построения ЭС. Этот инструмент отличается от обычных языков программирования тем, что обеспечивает удобные способы представления сложных высокоуровневых понятий.

Пользователь — это человек, который использует уже построенную ЭС. Так, пользователем может быть юрист, использующий её для квалификации конкретного случая; студент, которому ЭС помогает изучать информатику и т. д. Термин пользователь несколько неоднозначен. Обычно он обозначает конечного пользователя.

**Вопрос № 7. Этапы разработки экспертных систем.**

На этапе **идентификации** разработчик и эксперт определяют существенные особенности задачи: сама за­дача, участники про­цесса разработки, тре­буемые ресурсы, а также цели и задачи создания экспертной системы.

На этапе **концептуализации**инженер знаний и эксперт ре­шают, какие понятия, отношения и механизмы управления нужны для описания решения задач в избранной области. На этой стадии обычно требуется рассмотреть вопрос о гранулярности, т. е. какой сте­пенью подробности нужно представлять знания.

**Формализация**состоит в выражении ключевых понятий и отношений некоторым формальным способом, обычно в рамках схемы, диктуемой языком построения экспертных систем.

На этапе **реализации**разработчик превращает формализо­ванные знания в работающую компьютерную программу. Напи­сание программы требует содержания, формы и согласования. Содержание берется из предметных знаний, сделанных явными в ходе формализации, т. е. структур данных, правил вывода и стратегий управления, необходимых для решения задачи. Форма задается языком, выбранным для разработки системы. Согласование включает в себя комбинирование и реорганиза­цию различных порций знаний с целью устранить глобальные неувязки между спецификациями структур данных, правил и схем управления.

**Тестирование**включает оценивание качества ра­боты и полезности программы-прототипа и ее пересмотр, если это необходимо. Эксперт обычно оценивает прототип и помо­гает разработчику его пересмотреть. Как только прототип ис­пытан на нескольких примерах, его следует протестировать на многих задачах, чтобы оценить качество его работы и его полезность. При таком оценивании могут обнаружиться недо­статки схемы представления, например отсутствие нужных по­нятий или отношений, неправильный уровень детализации зна­ний или неудобные механизмы управления. Эти недостатки мо­гут заставить разработчиков заново пройти через разные этапы разработки, переформулировать понятия, уточнить правила вы­вода и пересмотреть схему управления.

**Прототипи́рование** — быстрая «черновая» реализация базовой функциональности для анализа работы системы в целом. На этапе прототипирования малыми усилиями создается работающая система Во время прототипирования видна более детальная картина устройства системы. Используется в машино - и [приборостроении](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%B8%D0%B1%D0%BE%D1%80%D0%BE%D1%81%D1%82%D1%80%D0%BE%D0%B5%D0%BD%D0%B8%D0%B5), [программировании](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5) и во многих других областях техники. Прототипирование, по мнению некоторых разработчиков, является самым важным этапом [разработки](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D0%B7%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%BA%D0%B0_%D0%BD%D0%BE%D0%B2%D0%BE%D0%B3%D0%BE_%D0%BF%D1%80%D0%BE%D0%B4%D1%83%D0%BA%D1%82%D0%B0).

# **Вопрос № 8. Средства построения экспертных систем.**

**Традиционные языки программирования**. В эту группу инструментальных средств входят традиционные языки программирования (С, C++, Basic, Pascal, Fortran и т.д.), ориентированные в основном на численные алгоритмы и слабо подходящие для работы с символьными и логическими данными. Поэтому создание систем искусственного интеллекта на основе этих языков требует большой работы программистов. Однако большим достоинством этих языков является высокая эффективность, связанная с их близостью к традиционной машинной архитектуре. Кроме того, использование традиционных языков программирования позволяет включать интеллектуальные подсистемы (например, интегрированные экспертные системы) в крупные программные комплексы общего назначения. Среди традиционных языков наиболее удобными считаются объектно-ориентированные (Pascal, C++). Это связано с тем, что парадигма объектно-ориентированного программирования тесно связана с фреймовой моделью представления знаний. Кроме того, традиционные языки программирования используются для создания других классов инструментальных средств искусственного интеллекта.

**Языки искусственного интеллекта.** Это прежде всего Лисп (LISP) и Пролог (Prolog) – наиболее распространённые языки, предназначенные для решения задач искусственного интеллекта. Универсальность этих языков меньшая, нежели традиционных языков, но её потерю языки искусственного интеллекта компенсируют богатыми возможностями по работе с символьными и логическими данными, что крайне важно для задач искусственного интеллекта. На основе языков искусственного интеллекта создаются специализированные компьютеры (например, Лисп-машины), предназначенные для решения задач искусственного интеллекта. Недостаток этих языков – неприменимость для создания гибридных экспертных систем.

**Специальный программный инструментарий.** В эту группу программных средств искусственного интеллекта входят специальные инструментарии общего назначения. Как правило, это библиотеки и надстройки над языком искусственного интеллекта Лисп: KEE (Knowledge Engineering Environment), FRL (Frame Representation Language), KRL (Knowledge Representation Language), ARTS и другие, позволяющие пользователям работать с заготовками экспертных систем на более высоком уровне, нежели это возможно в обычных языках искусственного интеллекта.

**«Оболочки».** Под «оболочками» (shells) понимают «пустые» версии существующих экспертных систем, т.е. готовые экспертные системы без базы знаний. Примером такой оболочки может служить EMYCIN (Empty MYCIN – пустой MYCIN), которая представляет собой незаполненную экспертную систему MYCIN. Достоинство оболочек в том, что они вообще не требуют работы программистов для создания готовой экспертной системы. Требуется только специалисты в предметной области для заполнения базы знаний. Однако если некоторая предметная область плохо укладывается в модель, используемую в некоторой оболочке, заполнить базу знаний в этом случае весьма не просто.

# Вопрос № 9. Методы тестирования экспертных систем.

Наиболее перспективный подход для оценки эффективности применения различных методов тестирования на разных этапах разработки ЭС строится на базе использования концепции мутаций программного продукта. Под мутацией или мутантом понимается единичная ошибка, допущенная в процессе проектирования.

|  |  |
| --- | --- |
| Методы тестирования экспертных систем | |
| Название метода | Описание |
| А. Тестирование на основе концепции "черного ящика" | Набор тестируемых ситуаций генерируется без учета используемых в системе методов решения задачи |
| Случайное тестирование | Тестируемые ситуации выбираются случайным образом из пространства входных наборов данных |
| Выборочное тестирование входов | Пространство входных наборов данных разбивается на выборки, для которых определяются ситуации для тестирования |
| Выборочное тестирование выходов | Тестируемые ситуации определяются на основе выборок, сформированных для выходных наборов данных |
| Б. Тестирование на основе концепции "белого ящика" | Тестируемые ситуации учитывают внутреннюю структуру системы в дополнение к входным и ожидаемым выходным наборам данных |
| Тестирование потоков данных | Анализ системы для выявления аномальных ситуаций, связанных с описанием, использованием и уничтожением переменных |
| Тестирование динамических потоков | Тестируемые ситуации генерируются для прохождения различных ветвей исполнения программы |
| Тестирование причин и следствий | Причины и следствия определяются на основе анализа решений, и тестовые ситуации формируются путем комбинации причин |
| В. Тестирование полноты базы знаний | Тестирование правил на внутреннюю неполноту |
| Поиск конфликтных правил | Поиск правил, возбуждающихся в сходных ситуациях, но приводящих к различным результатам |
| Поиск избыточных правил | Поиск правил, возбуждающихся в сходных ситуациях и приводящих к сходным результатам |
| Поиск пересекающихся правил | Анализ системы на наличие правил, являющихся подмножеством других правил в части антецедентов или консеквентов, но не одновременно |
| Г. Тестирование целостности базы знаний | Тестирование правил на внутреннюю целостность |
| Поиск пропущенных правил | Анализ системы для нахождения пропущенных правил, приводящих к требуемым результатам |
| Поиск атрибутов без ссылок | Анализ системы для нахождения атрибутов, на которые не существует ссылок ни в одном правиле |
| Поиск атрибутов с некорректными значениями | Нахождение некорректных значений атрибутов, на которые ссылаются правила системы |

**10. Соотнесение понятий: информация, данные, знания и контент (определение и взаимосвязь). Классификация знаний**

**Знания**  являются низшим уровнем иерархии и определяются как часть знаний, критерий истинности которых субъективен у различных участников познавательного процесса.

Под **данными** понимаются неупорядоченные наблюдения, числа, слова, звуки, изображения. Это набор дискретных, объективных фактов.

Например, структурированные записи о трансакциях (сколько, когда и кто купил, сколько и когда заплатил и пр.) являются данными. Эти данные не говорят о том, почему покупатель пришел именно сюда и придет ли он еще раз[[2]](https://studme.org/53195/politekonomiya/ponyatie_znaniy_sootnoshenie_ponyatiy_znanie_informatsiya_dannye" \l "gads_btm).

Когда данные организованы, упорядочены, сгруппированы и категоризированы, они становятся информацией.

**Информация** – это совокупность данных, упорядоченная с определенной целью, придающей им смысл.

Говоря о знаниях, необходимых сотрудникам компании, нельзя отождествлять их с понятием «информация». Если информация состоит из данных, организованных определенным образом для того, чтобы охарактеризовать определенную ситуацию, условия или явление, то знание состоит из фактов, ожиданий и концепций, ментальных моделей, истин и убеждений, ожиданий и ноу-хау. Информация используется для описания и определения различных предметов и ситуаций, в которых они развивались и существуют, получается в результате применения знаний к описанию и объяснению явлений на основе полученных данных, но это еще не делает ее знанием. Знания же применяются для оценки и управления ситуацией, [принятия решений](https://www.e-xecutive.ru/wiki/index.php/%D0%9F%D1%80%D0%B8%D0%BD%D1%8F%D1%82%D0%B8%D0%B5_%D1%80%D0%B5%D1%88%D0%B5%D0%BD%D0%B8%D1%8F), планирования и действия.

Классификация знаний составляется на основе таксоно́мии (от древне-греческого Τάξις - строй, порядок и νόμος – закон). Это учение о принципах и практике классификации и систематизации.

Классификация знаний представляет собой довольно сложную систему, где все знания изначально классифицируются по признакам «ясность», «доступность», «пропозициональность» и «уровень абстракции».

Далее ясные знания подразделяются на явные (сформулированные и обобщенные) и неявные, которые отражаются в действиях, опыте, применительно к конкретной ситуации.

По признаку доступности выделяются знания индивидуальные (созданные и используемые конкретным человеком) и коллективные (групповые, организационные, межорганизационные), которые создаются и используются в коллективной деятельности определенной группы людей.

Наличие категории «уровень абстракции» позволяет выделять знания конкретные, специализированные, которые дают представление о конкретном предмете или явлении, способе действия в ситуации, и общее, которые объясняют роль, место и взаимосвязь объекта с другими объектами.

Для практической деятельности наибольшую ценность представляют знания, которые классифицируются как пропозициональные. Они включают в себя такие типы знаний, как декларативные (знания о чем-то), процедурные (знание осуществления процесса), каузальные, причинно-следственные (знание причин явлений), знания условий и времени выполнения действия, знания отношений объекта действий с другими объектами.

11. Инженерия знания. Структура инженерии знаний. Семиотика и её составляющие. Пирамида знаний   
Инженер знаний- это специалист по искуственному интеллекуту, проектирующий и создающий Экспертную систему илди другую информациооную систему

Инженер знаний выполняет важные функции при разработке БЗ. Он должен хорошо ориентироваться в проблемной области и быть неплохим психологом, чтобы общаться с экспертом в процессе приобретения знаний. Вместе с тем он должен хорошо знать и возможности программного обеспечения компьютеров, чтобы структурировать знания для хранения и работы с ними.  Основным источником знаний о проблемной области является человек-эксперт. Экс перт - специалист, который за годы обучения и практической деятельности научился эффективно решать задачи, относящиеся к конкретной предметной области.

Инженер знаний работает с ним в режиме диалога или интервью и формирует необходимый объем знаний и сведений для работы с объектом. Возможно также использование *опросников,* которые затем соответствующим образом обрабатываются.

|  |  |
| --- | --- |
| 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. |

В программу симпозиума входили следующие секции: 1) естественный язык как знаковая система; 2) знаковые системы письма и дешифровка; 3) неязыковые системы коммуникации; 4) искусственные языки; 5) моделирующие семиотические системы; 6) искусство как семиотическая система; 7) структурное и математическое изучение литературных произведений

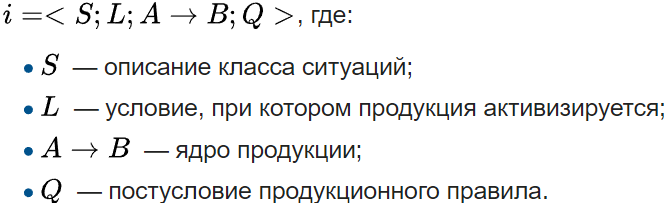
12. Модели представления знания. Продукционные модели

**Продукционная модель знания** — [модель](https://ru.wikipedia.org/wiki/%D0%9C%D0%BE%D0%B4%D0%B5%D0%BB%D1%8C), основанная на [правилах](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%B0%D0%B2%D0%B8%D0%BB%D0%B0), позволяет представить знание в виде предложений типа «Если (условие), то (действие)».

Продукционная модель — фрагменты [Семантической сети](https://ru.wikipedia.org/wiki/%D0%A1%D0%B5%D0%BC%D0%B0%D0%BD%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D0%B5%D1%82%D1%8C), основанные на временных отношениях между состояниями объектов.

Продукционная модель обладает тем недостатком, что при накоплении достаточно большого числа (порядка нескольких сотен) продукций они начинают вследствие необратимости дизъюнкций противоречить друг другу. В этом случае разработчики начинают усложнять систему, включая в неё модули нечёткого вывода или иные средства разрешения конфликтов, — правила по приоритету, правила по глубине, эвристические механизмы исключений, возврата и т. п.

В общем случае продукционную модель можно представить в следующем виде:



**13. Модели представления знания. Семантические сети**

*(-- начало воды для общего понимания)*

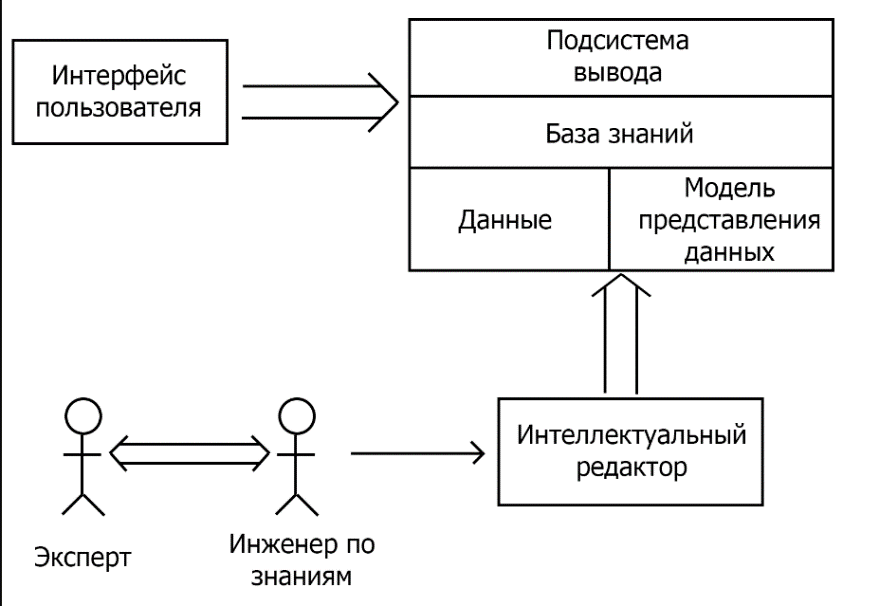
*Инженерия знаний* (англ. knowledge engineering) — область наук об искусственном интеллекте, связанная с разработкой экспертных систем и баз знаний. Изучает методы и средства извлечения, представления, структурирования и использования знаний.

Разработка экспертных систем (ЭС) неразрывно связано с использованием методов инженерии знаний и, нередко, между ними ставят знак равенства.

*Экспе́ртная систе́ма* (ЭС, англ. expert system) — компьютерная система, способная частично заменить специалиста-эксперта в разрешении проблемной ситуации.

*Экспертная система* (система, основанная на знаниях) – это программа, аккумулирующая знания специалистов (экспертов) в определенной предметной области и оперирующая знаниями с целью выработки рекомендаций или решения проблем.

*Экспертная система* (ЭС) — это информационная система, назначение которой частично или полностью заменить эксперта в той или иной предметной области. Подобные интеллектуальные системы эффективно применяются в таких областях, как логистика, управление воздушными полетами, управление театром военных действий. Основною направленной деятельностью предсказание, прогнозирование в рамках определенного аспекта в предметной области.



*(-- конец воды)*

*Знания* — это правила, законы, закономерности получены в результате профессиональной деятельности в пределах предметной области.

*Модель представления знаний (МПЗ)* — это способ задания знаний для хранения, удобного доступа и взаимодействия с ними, который подходит под задачу интеллектуальной системы. Выделяют: *продукционную МПЗ, семантическую МПЗ, фреймовую МПЗ, формально-логическую МПЗ.*

*Семантическая сеть* — информационная модель предметной области, имеющая вид ориентированного графа, вершины которого соответствуют объектам предметной области, а дуги (рёбра) задают отношения между ними. Объектами могут быть понятия, события, свойства, процессы. Таким образом, семантическая сеть является одним из способов представления знаний. В семантической сети роль вершин выполняют понятия базы знаний, а дуги (причем направленные) задают отношения между ними. Таким образом, семантическая сеть отражает семантику предметной области в виде понятий и отношений.

Ориентированный граф. Вершины графа — понятия, дуги — отношения между понятиями.

Классификация. *По количеству типов* *отношений*: *однородные* и *неоднородные* семантические сети. Однородные имею один тип отношения между всеми понятиями, следовательно, не однородные имею множество типов отношений. *По* *арности*: сети с *бинарными отношениями* (связывающими ровно два понятия). Бинарные отношения очень просты и удобно изображаются на графе в виде стрелки между двух концептов. *N-арные сети* (связывающие более двух объектов). *По размеру:* *Для решения конкретных задач*, например, тех которые решают системы искусственного интеллекта. Семантическая сеть *отраслевого масштаба* должна служить базой для создания конкретных систем, не претендуя на всеобщее значение. *Глобальная семантическая сеть*. Теоретически такая сеть должна существовать, поскольку всё в мире взаимосвязано. Возможно когда-нибудь такой сетью станет Всемирная паутина.

Типы отношений:

Иерархические:

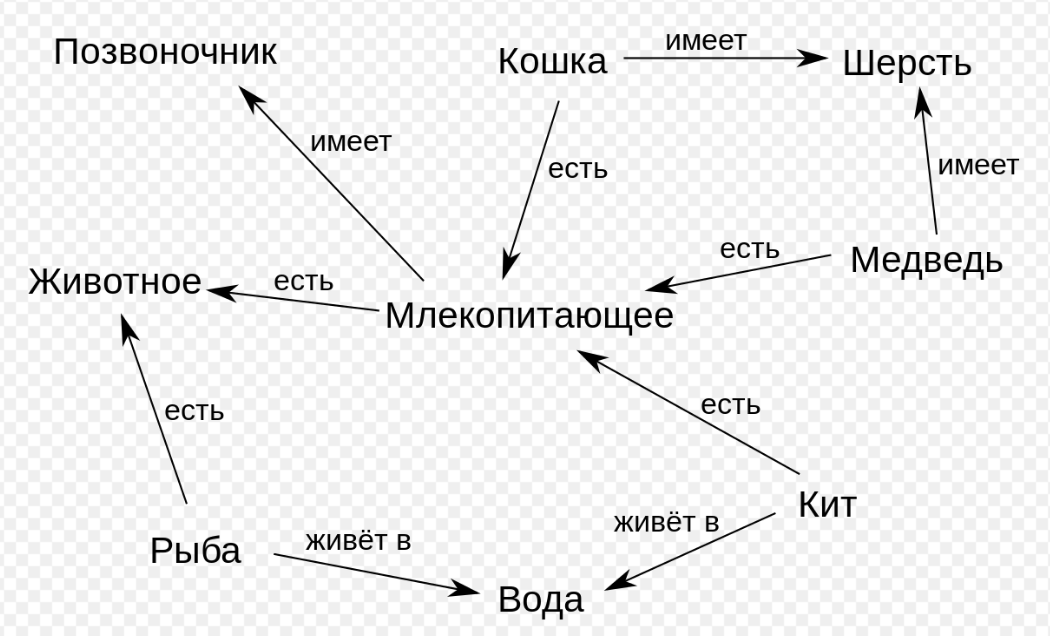
* часть — целое
* класс — подкласс

Вспомогательные:

* элемент — количество (больше, меньше, равно)
* атрибутивные (иметь свойство, иметь значение)
* логические (И, ИЛИ, НЕ)
* лингвистические

Недостаток: сложность извлечения знаний (обход графа).

Применение: WordNet — это электронный тезаурус/семантическая сеть для английского языка.



**14. Модели представления знаний. Фреймы**

*Знания* — это правила, законы, закономерности получены в результате профессиональной деятельности в пределах предметной области.

*Модель представления знаний (МПЗ)* — это способ задания знаний для хранения, удобного доступа и взаимодействия с ними, который подходит под задачу интеллектуальной системы. Выделяют: *продукционную МПЗ, семантическую МПЗ, фреймовую МПЗ, формально-логическую МПЗ.*

*Фрейм* (англ. frame — «каркас» или «рамка») — способ представления знаний в искусственном интеллекте, представляющий собой схему действий в реальной ситуации. Фрейм — это модель абстрактного образа, минимально возможное описание сущности какого-либо объекта, явления, события, ситуации, процесса.

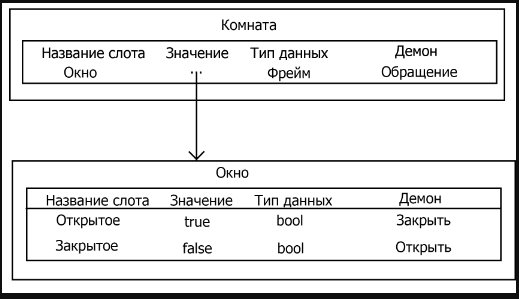
Различают фреймы-образцы, фреймы-экземпляры, фреймы-структуры, фреймы-роли, фреймы-сценарии, фреймы-ситуации. Система связанных фреймов может образовывать семантическую сеть. Применяются фреймы в ЭС и других интеллектуальных системах различного назначения.

Фрейм включает определённое знание по умолчанию, которое называется презумпцией (предположение, которое считается истинным, пока не доказана его ложность).

Фрейм — это образ, рамка, шаблон, которая описывает объект предметной области, с помощью слотов. *Слот* — это атрибут объекта. Слот имеет имя, значение, тип хранимых данных, демон. *Демон* — процедура, автоматически выполняющаяся при определенных условиях. Имя фрейма должно быть уникальным в пределах одной фреймовой модели. Имя слота должно быть уникальным в пределах одного фрейма.



Слот может хранить другой фрейм, тогда фреймовая модель вырождается в *сеть фреймов*.



**15. Психологический аспект инженерии знаний (контактный слой, процедурный слой, когнитивный слой)**

*Инженерия знаний* (англ. knowledge engineering) — область наук об искусственном интеллекте, связанная с разработкой экспертных систем и баз знаний. Изучает методы и средства извлечения, представления, структурирования и использования знаний.

*Извлечение знаний* – это процедура взаимодействия эксперта с источником знаний, в результате которой становятся явными процесс рассуждений специалистов при принятии решения и структура их представлений о предметной области.

Тот факт, что знания извлекаются в процессе общения двух или более специалистов, включающих инженера по знаниям и экспертов, обуславливает существование трех основных аспектов процесса извлечения знаний: *психологического, лингвистического, гносеологического.*

*Психологический аспект* - является наиболее важным, поскольку определяет успешность и эффективность общения инженера по знаниям и эксперта. Модель общения при извлечении знаний включает следующие структурные компоненты: участников общения (партнеров) - определяют *контактный слой* общения; средства общения (процедура) – определяют *процедурный слой*; предмет общения (знания) – определяют *когнитивный слой.*

*Контактный слой.* Следует обеспечить дружескую атмосферу для общения инженера по знаниям и эксперта. Для этого следует учитывать: пол и возраст, характеристики личности и темперамента (рекомендуемые составляющие для инженера по знаниям: доброжелательность, аналитичность, хорошая память, внимание, наблюдательность, воображение, впечатлительность, большая собранность, настойчивость, общительность, находчивость, сангвинический и холерический темперамент), мотивацию участников.

*Процедурный слой.* Параметры этого слоя описывают процесс проведения процедуры извлечения знаний – место, время, продолжительность и стиль общения. Одно из основных средств увеличения эффективности процесса извлечения знаний - использование наглядного материала. Для протоколирования результатов беседы используются следующие способы: запись на бумагу непосредственно по ходу беседы; запись на диктофон,; запоминание с последующей записью после собеседования. Большая часть информации поступает к инженеру по знаниям в форме предложений на естественном языке. Важны невербальные средства, такие, как интонация, мимика, жесты. Опытный инженер по знаниям старается по возможности записывать и эту дополнительную интонацию.

*Когнитивный слой*. Для эффективного моделирования когнитивных (познавательных) процессов необходимо учитывать их природу. Под когнитивным стилем человека понимается совокупность критериев предпочтения при решении задач и познании мира, специфическая для каждого индивида. Когнитивный стиль определяет не столько эффективность деятельности, сколько способ достижения результата. Характеристики когнитивных стилей:

* *Полезависимость - поленезависимость*. Поленезависимость позволяет человеку акцентировать внимание лишь на тех аспектах проблемы, которые необходимы для решения конкретной задачи, и отбрасывать все лишнее. Это желательный фактор и для эксперта, и для инженера по знаниям. Полезависимые люди лучше общаются, они более контактны. Для общения особенно удачны гетерогенные (смешанные) пары: «полезависимый - поленезависимый».
* *Импульсивность – рефлексивность*. Под импульсивностью понимается быстрое принятие решения (часто без достаточного обоснования), а под рефлексивностью - склонность к рассудительности. Рефлексивность связана со способностью к формированию понятий и решению логических задач. Таким образом, и инженеру по знаниям, и эксперту желательно быть рефлексивными, хотя собственный стиль изменяется лишь частично и с большим напряжением.
* *Ригидность - гибкость.* - способность человека изменять установки и свою точку зрения в соответствии с изменяющейся ситуацией. Ригидность допустима у экспертов, недопустима у инженеров по знаниям.
* *Когнитивная эквивалентность* - способность человека к различению понятий и разбиению их на классы и подклассы. Чем уже диапазон когнитивной эквивалентности, тем более тонкую классификацию способен провести индивид, тем большее количество признаков понятий он может выделить.

**16. Лингвистический аспект инженерии знаний («общий код», понятийная структура, словарь пользователя)**

*Инженерия знаний* (англ. knowledge engineering) — область наук об искусственном интеллекте, связанная с разработкой экспертных систем и баз знаний. Изучает методы и средства извлечения, представления, структурирования и использования знаний.

*Извлечение знаний* – это процедура взаимодействия эксперта с источником знаний, в результате которой становятся явными процесс рассуждений специалистов при принятии решения и структура их представлений о предметной области.

Тот факт, что знания извлекаются в процессе общения двух или более специалистов, включающих инженера по знаниям и экспертов, обуславливает существование трех основных аспектов процесса извлечения знаний: *психологического, лингвистического, гносеологического.*

*Лингвистический аспект* - касается исследований языковых проблем, поскольку язык — основное средство общения в процессе извлечения знаний. Инженер по знаниям, приступивший к процессу приобретения информации, сталкивается с тремя слоями лингвистических проблем: формирование *«общего кода»;* построение *понятийной структуры*; создание *словаря пользователя.*

*Общий код* как средство общения устраняет проблему языковых ножниц между профессиональной терминологией экспертов и инженера по знаниям. Язык - это основное средство мышления наряду с другими знаковыми системами "внутреннего пользования". Языки, на которых говорят и размышляют аналитик и эксперт, могут существенно отличаться. Есть два языка - *язык аналитика* (термины предметной области, которые он почерпнул из специальной литературы в период подготовки; общенаучная терминология из его "теоретического багажа"; бытовой разговорный язык, которым пользуется аналитик) и *язык эксперта* (специальная терминология, принятая в предметной области; общенаучная терминология; бытовой язык; неологизмы, созданных экспертом за время работы (его профессиональный жаргон)). Если считать, что бытовой и общенаучный языки у двух участников общения примерно совпадают, то некоторый общий язык, или код, который необходимо выработать партнерам для успешного взаимодействия, будет складываться из потоков их языков. В дальнейшем этот общий код преобразуется в некоторую понятийную (семантическую) сеть, которая является прообразом поля знаний предметной области. Выработка общего кода начинается с выписывания инженером по знаниям всех терминов, употребляемых экспертом, и уточнения их смысла — фактически это составление словаря предметной области. Затем следует группирование терминов и выбор синонимов. Разработка общего кода заканчивается составлением словаря терминов предметной области с предварительной группировкой их по смыслу, то есть по понятийной близости (это уже первый шаг структурирования знаний)



*Понятийная структура*. Основная особенность естественного интеллекта и памяти в частности - это связанность всех понятий в некоторую сеть. Поэтому для разработки базы знаний нужен не словарь, а энциклопедия, в которой все термины объяснены в словарных статьях со ссылками на другие термины. Таким образом, лингвистическая работа инженера по знаниям на данном слое проблем заключается в построении таких связанных фрагментов с помощью "сшивания" терминов. При тщательной работе аналитика и эксперта в понятийных структурах начинает проглядывать иерархия понятий, что в общем согласуется с результатами когнитивной психологии. *Иерархия понятий* - это глобальная схема, которая может быть в основе концептуального анализа структуры знаний любой предметной области. Работа по составлению словаря и понятийной структуры требует лингвистического "чутья", легкости манипулирования терминами и богатого словарного запаса инженера по знаниям.

*Словарь пользователя.* Лингвистические результаты, соотнесенные со слоями общего кода и понятийной структуры, направлены на создание адекватной базы знаний. Однако не следует забывать, что профессиональный уровень конечного пользователя может не позволить ему применить специальный язык предметной области в полном объеме. Для разработки пользовательского интерфейса необходима дополнительная доработка словаря общего кода с поправкой на доступность и "прозрачность" системы. Общий код и язык пользователя различаются, т.к. различаются знания эксперта и пользователя.

**17. Методологический (гносеологический) аспект инженерии знаний. Гносеологическая цепочка. Историзм. Внутренняя согласованность. Системность. Объективность.**

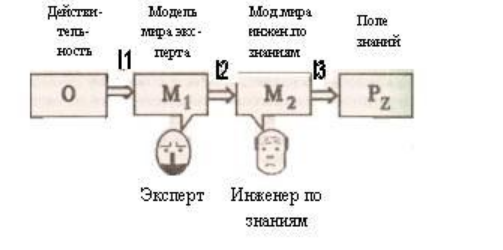
*Инженерия знаний* (англ. knowledge engineering) — область наук об искусственном интеллекте, связанная с разработкой экспертных систем и баз знаний. Изучает методы и средства извлечения, представления, структурирования и использования знаний.

*Извлечение знаний* – это процедура взаимодействия эксперта с источником знаний, в результате которой становятся явными процесс рассуждений специалистов при принятии решения и структура их представлений о предметной области.

Тот факт, что знания извлекаются в процессе общения двух или более специалистов, включающих инженера по знаниям и экспертов, обуславливает существование трех основных аспектов процесса извлечения знаний: *психологического, лингвистического, гносеологического.*

Гносеология - это раздел философии, связанный с теорией познания, или теорией отражения действительности в сознании человека. Гносеологический аспект извлечения знаний объединяет методологические проблемы получения нового научного знания, поскольку при создании БЗ эксперт часто впервые формулирует некоторые закономерности, до того момента составлявшие его личный опыт.

Инженерия знаний как наука гносеологична - действительность (О) сначала отражается в сознании эксперта (М1), а затем деятельность и опыт эксперта интерпретируются сознанием инженера по знаниям (M2), что служит уже основой для построения третьей интерпретации (Pz) - поля знаний экспертной системы. Процесс познания в сущности направлен на создание внутреннего представления окружающего мира в сознании человека.



*Гносеологическая цепочка:* факт -> обобщённый факт -> эмпирический закон -> теоретический закон.

Методологические критерии: Гносеологический аспект A3 = {S31, S32, S33, S34} = {внутренняя согласованность, системность, объективность, историзм}.

*Историзм (S34)* связан с развитием. Познание настоящего — есть познание породившего его прошлого. И хотя большинство экспертных систем дают «горизонтальный» срез знаний — без учета времени (в статике), инженер по знаниям должен всегда рассматривать процессы с учетом временных изменений — как связь с прошлым, так и связь с будущим.

*Внутренняя согласованность (S31).* S31 = {s31\_i} = {модальность, противоречивость, неполнота}. Модальность знания - возможность его существования в различных категориях, то есть в конструкциях существования и долженствования. Кроме того, приходится различать такие оттенки модальности, как (знает, думает, хочет, считает). Возможная противоречивость эмпирического знания — естественное следствие из основных законов диалектики, и противоречия эти не всегда должны разрешаться в поле знаний, а напротив, именно противоречия служат чаще всего отправной точкой в рассуждениях экспертов. Неполнота знания связана с невозможностью полного описания предметной области. Задача аналитика эту неполноту ограничить определенными рамками «полноты», то есть сузить границы предметной области либо ввести ряд ограничений и допущений, упрощающих проблему.

*Системность (S32).* - рассмотрение любой предметной области с позиций закономерностей системного целого и взаимодействия составляющих его частей.

*Объективность (S33).* Процесс познания глубоко субъективен, то есть он существенно зависит от особенностей самого познающего субъекта. Таким образом, субъективность начинается уже с описания фактов и увеличивается по мере углубления идеализации объектов.Следовательно, более корректно говорить о глубине понимания, чем об объективности знания. Понимание — это сотворчество, процесс истолкования объекта с точки зрения субъекта. Это сложный и неоднозначный процесс, совершающийся в глубинах человеческого сознания и требующий мобилизации всех интеллектуальных и эмоциональных способностей человека. Все свои усилия аналитик должен сосредоточить на понимании проблемы.

**18. Текстологические методы извлечения знаний (анализ учебников, анализ литературы, анализ документов). Алгоритм извлечения знаний из текста**

*Инженерия знаний* (англ. knowledge engineering) — область наук об искусственном интеллекте, связанная с разработкой экспертных систем и баз знаний. Изучает методы и средства извлечения, представления, структурирования и использования знаний.

*Извлечение знаний* – это процедура взаимодействия эксперта с источником знаний, в результате которой становятся явными процесс рассуждений специалистов при принятии решения и структура их представлений о предметной области.

Классификация методов извлечения знаний выполняется в зависимости от источника знаний:

* *коммуникативные* методы охватывают способы и процедуры контактов инженера по знаниям с непосредственным источником знаний - экспертами того или иного производства (предметной области). Применяются для слабо/плохо структурированных и слабо/плохо документированных предметных областей;
* *текстологические* включают методы извлечения знаний из документов (методик, регламентов, служебных руководств) и специальной литературы (статей, монографий, учебников). Применяются для хорошо структурированных и хорошо документированных предметных областей.

При извлечении знаний в качестве *текстологических* источников информации выступают:

* *учебники*. Наиболее доступные для понимания источники, поскольку логика изложения обычно соответствует логике предмета, и макроструктура такого текста будет более значима, чем структура текста специальной статьи, Эффективны при предварительном ознакомлении с предметной областью и при более углубленном понимании других источников текстологических знаний;
* *специальная и научная литература*(статьи, монографии). Предполагают предварительную специальную подготовку, потому могут использоваться на последующих этапах извлечения знаний;
* *документы*. Характеризуются сжатостью изложения и практическим отсутствием комментариев, т.е. фоновых знаний, облегчающих понимание для неспециалиста. Включают большое количество имплицитной информации, потому могут эффективно использоваться на последующих этапах извлечения знаний.

*Алгоритм извлечения знаний из текста:*

1. Составление базового списка литературы для ознакомления с предметной областью и чтение по списку;
2. Выбор текста для извлечения знаний;
3. Первое знакомство с текстом (беглое прочтение), для определения значения незнакомых слов – консультации со специалистами или привлечение справочной литературы;
4. Формирование первой гипотезы о макроструктуре текста.
5. Внимательное прочтение текста с выписыванием ключевых слов и выражений, т.е. выделение «смысловых вех» (компрессия текста).
6. Определение связей между ключевыми словами, разработка макроструктуры текста в форме графа или сжатого текста (реферата).
7. Формирование поля знаний на основании макроструктуры текста.

Преимущества текстологических методов заключаются в отсутствии необходимости для инженера по знаниям в общении с экспертом. Однако уповать только на текстологические источники знаний принципиально неверно, поскольку лишь эксперты обладают уникальным опытом, интуицией, наработанной годами профессионального труда, которые принципиально не могут быть представлены в текстовом виде. Поэтому на практике требуется разумное совмещение обоих подходов (коммуникативный + текстологический) к извлечению знаний.

**19. Коммуникативные методы извлечения знаний: пассивные методы (наблюдение, протокол «мыслей вслух», лекции)**

Коммуникативные методы – это набор приемов и процедур, предпочитающих контакт когнитолога с непосредственным источником знаний – экспертом, а текстологические включают методы извлечения знаний из документов и специальной литературы.

В свою очередь коммуникативные методы делятся на пассивные и активные. Пассивныеметоды извлечения знаний включают такие методы, где ведущая роль фактически передается эксперту, а когнитолог только фиксирует рассуждения эксперта во время работы по принятию решений. К этим методам относятся наблюдения, анализ протоколов «мыслей вслух» и лекции.

В процессе наблюденийкогнитолог находится рядом с экспертом во время его профессиональной деятельности и протоколирует все его действия. Эксперт максимально комментирует свои действия. Непременное условие этого метода – невмешательство когнитолога в работу эксперта. Наблюдения – один из распространенных методов извлечения знаний на начальных этапах разработки ЭС.

Протоколирование «мыслей вслух»отличается от наблюдений тем, что эксперт не просто комментирует, но и объясняет цепочку рассуждений при решении конкретной задачи. Основной трудностью при протоколировании «мыслей вслух» является принципиальная сложность для любого человека объяснить, как он думает, поскольку люди не всегда в состоянии достоверно описать мыслительные процессы. Кроме того, часть знаний, хранящихся в невербальной форме, слабо коррелируют с их словесным описанием (например, различные процедурные знания типа «Как завязать галстук»). Обычно «мысли вслух» дополняются одним из активных методов для реализации обратной связи с экспертом.

Лекция является самым старым способом передачи знаний. Использование лекции для извлечения знаний целесообразно, если эксперт обладает лекторским мастерством. В процессе лекции когнитолог может задавать уточняющие вопросы. Как и другие пассивные методы, лекции используются для извлечения знаний на начальных стадиях разработки ЭС.

**20. Коммуникативные методы извлечения знаний: активные групповые методы («мозговой штурм», круглый стол, ролевые игры)**

К активным групповым методам извлечения знаний относятся «мозговой штурм», дискуссии за круглым столом и ролевые игры.

«Мозговой штурм»,или «мозговая атака» – один из наиболее распространенных методов раскрепощения и активизации творческого мышления. Основная идея «штурма» – это отделение процедуры генерирования идей в замкнутой группе экспертов от процесса анализа и оценки высказанных идей. Как правило, «штурм» длится не более 40 минут с числом участников до 10 человек. Экспертам предлагается высказывать любые идеи на заданную тему (критика запрещена). Оценивает результаты группа экспертов, не участвовавшая в «штурме».

Метод круглого столапредусматривает обсуждение какой-либо проблемы из выбранной предметной области, в котором принимают участие с равными правами несколько экспертов. Задача дискуссии – коллективно, с разных точек зрения, под разными углами исследовать спорные гипотезы предметной области. По ходу дискуссии важно проследить, чтобы слишком эмоциональные и разговорчивые эксперты не подменили тему и чтобы критика позиций друг друга была обоснованной.

Ролевые игры предусматривают участие в экспертной игре нескольких экспертов. Роль – комплекс образцов поведения. Когнитолог является режиссером и сценаристом, и ему предоставляется полная свобода в выборе формы проведения игры. Из трех основных типов деловых игр (учебных, планово-производственных и исследовательских) к экспертам ближе всего исследовательские, которые используются для анализа систем, проверки правил принятия решений.

**21. Коммуникативные методы извлечения знаний: активные индивидуальные методы (анкетирование, интервью, экспертные игры)**

В **активных**методах извлечения знаний ведущая роль принадлежит когнитологу. Активные методы подразделяются на индивидуальные и групповые.**Активные индивидуальные**методы извлечения знаний в настоящее время являются наиболее распространенными и включают анкетирование, интервью, свободный диалог, игры с экспертом.

**Анкетирование –**наиболее жесткий метод (стандартизированный). Когнитолог заранее составляет вопросник и осуществляет опрос нескольких экспертов. Существует несколько общих рекомендаций при составлении анкет:

* анкета не должна быть монотонной и однообразной;
* анкета не должна быть приспособлена к языку экспертов;
* поскольку вопросы влияют друг на друга, их последовательность должна быть строго продумана;
* желательно стремиться к оптимальной избыточности;
* язык анкеты должен быть ясным, понятным, предельно вежливым.

Под **интервью**понимается специфическая форма общения когнитолога с экспертом, в которой когнитолог задает эксперту серию заранее подготовленных вопросов. В интервью когнитолог имеет возможность в зависимости от ситуации ряд вопросов опускать или вставлять новые вопросы, изменять темп, разнообразить ситуацию общения. По форме вопросы подразделяются на открытые (для ответа предоставляют полную свободу) и закрытые (предполагают фиксированные ответы), личные и безличные, прямые и косвенные, вербальные и с использованием наглядного материала. По функции – на основные, зондирующие и контрольные, по воздействию вопросы подразделяются на нейтральные и наводящие.

**Свободный диалог**– это метод извлечения знаний в форме беседы когнитолога и эксперта. Диалог предполагает выбор правильного темпа и ритма беседы, поскольку при больших паузах эксперт отвлекается, а при высоком темпе быстро утомляются оба собеседника.

**Игра с экспертом**предполагает, что эксперт и когнитолог в моделируемой ситуации берут на себя некоторые роли, например «учитель» и «ученик». В процессе игры эксперт поправляет ошибки «ученика». Существуют следующие советы когнитологу по проведению индивидуальных игр: играть смело, нешаблонно; не навязывать игру эксперту, если он не расположен; в игре «не давить на эксперта» и не забывать цели игры.

**22. Стратегии получения знаний (извлечение знаний, приобретение знаний, формирование знаний). Структура процесса общения.**

**Методологические и технологические проблемы получения знаний**

При формировании поля знаний ключевым вопросом является сам процесс получения знаний, когда происходит перенос компетентности экспертов на инженеров по знаниям. Для названия этого процесса в литературе по ЭС получило распространение несколько терминов: приобретение, добыча, извлечение, получение, выявление, формирование знаний.

Термин «приобретение» трактуется либо очень широко — тогда он включает весь процесс передачи знаний от эксперта к базе знаний ЭС, либо уже как способ автоматизированного построения базы знаний посредством диалога эксперта и специальной программы (при этом структура поля знаний заранее закладывается в программу). В обоих случаях термин «приобретение» не касается самого таинства экстрагирования структуры знаний из потока информации о предметной области. Этот процесс описывается понятием «извлечение».

Извлечение знаний (knowledge elicitation) — это процедура взаимодействия эксперта с источником знаний, в результате которой становятся явными процесс рассуждений специалистов при принятии решения и структура их представлений о предметной области.

Формирование знаний (machine learning) — процесс анализа данных и выявление скрытых закономерностей с использованием специального математического аппарата и программных средств.

В структуре общения различают:

1) коммуникативную сторону;

2) интерактивную сторону;

3) перцептивную сторону.

Под коммуникативной стороной общения в узком смысле слова понимается процесс обмена информацией между людьми. Однако следует понимать, что, в отличие от простого «движения информации», между двумя устройствами мы имеем дело с отношением двоих индивидов, говорящих «на одном языке», являющихся активными субъектами и влияющими друг на друга в процессе общения.

Интерактивная сторона общения — условный термин, обозначающий характеристику компонентов общения, связанных со взаимодействием людей, непосредственной организацией их совместной деятельности. Цели общения отражают потребности совместной деятельности людей. Общение всегда должно предполагать некоторый результат — изменение поведения и деятельности других людей.

Перцептивная сторона общения: в процессе общения должно присутствовать взаимопонимание между участниками этого процесса. Само взаимопонимание может быть здесь истолковано по-разному: или как понимание целей, мотивов, установок партнера по взаимодействию, или как не только понимание, но и принятие, разделение этих целей, мотивов, установок.

**23. Методы структурирования знаний и данных и их классификация. Объектно-структурный анализ. Стратификация знаний предметной области.**

Структурирование знаний – процесс анализа знаний данной предметной области и синтеза в структуры,не зависящие от программной реализации

Методология структурирования знаний предполагает следующие обязательные элементы:

1. Абстрагирование ( выявление наиболее важных сущностей предметной области)

2. Иерархия ( взаимозависимость абстракций)

3. Выделение типов ( классов абстракций с условием наследования)

4. Модульность ( разбиение области знаний на отдельные модули)

5. Легкость в восприятии

Системность ( четкая взаимосвязь между понятиями)

7. Формализация ( создание формального языка (правил)оперирования с абстракциями)

Можно предложить в качестве базисной методологии структурного анализа знаний и формирования поля знаний обобщенный объектно-струк­турный подход (ОСП), последовательно разработанный от математического обо­снования до технологии и программной реализации.

Основные постулаты этого подхода заимствованы из ООП и расширены.

1. Системность (взаимосвязь между понятиями).

2. Абстрагирование (выявление существенных характеристик понятия, кото­рые отличают его от других).

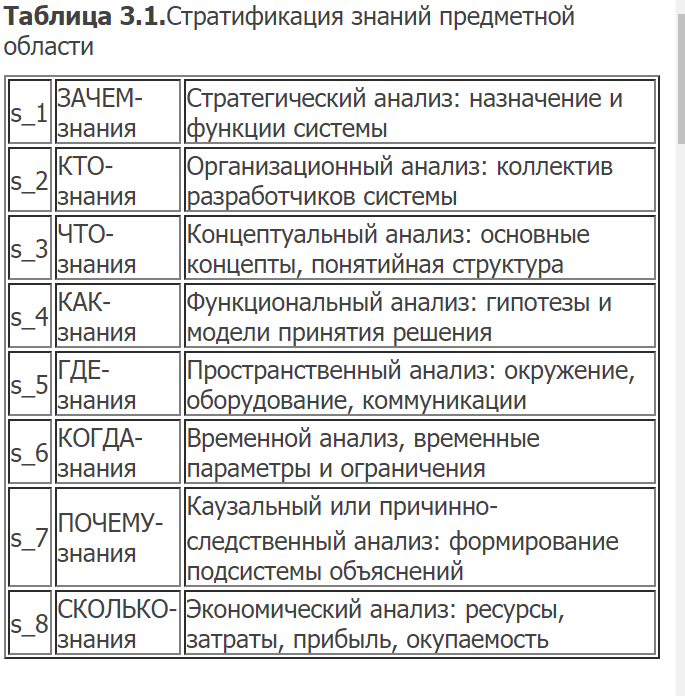
3. Иерархия (ранжирование на упорядоченные системы абстракций).

4. Типизация (выделение классов понятий с частичным наследованием свойств в подклассах).

5. Модульность (разбиение задачи на подзадачи или «возможные миры»).

6. Наглядность и простота нотации.

Объектно-структурный подход подразумевает интегрированное использование сформулированных выше постулатов от первой до последней стадий разработки БЗ интеллектуальных и обучающих систем. На основе ОСП предлагается алго­ритм объектно-структурного анализа (ОСА) предметной области, позволяюще­го оптимизировать и упорядочить достаточно размытые процедуры структури­рования знаний.



Объектно-структурный анализ подразумевает разработку и использование мат­рицы ОСА (см. табл. 3.2), которая позволяет всю собранную информацию деза­грегировать последовательно по слоям-стратам (вертикальный анализ), а затем по уровням - от уровня проблемы до уровня подзадачи (горизонтальный ана­лиз). Или наоборот - сначала по уровням, а потом по стратам.

**24. Методы структурирования знаний и данных и их классификация. Визуальные методы. Стадии структурирования знаний**

(Смотри синий абзац выше и будет тебе счастье, детка)

Одним из способов анализа получаемой информации является ее структурирование. Структурирование есть ни что иное, как расположение в определенном порядке, или по определенной схеме.

Визуализация информации – представление информации в виде графиков, диаграмм, структурных схем, таблиц, карт и т.д.

Теоретическую основу методов визуализации информации и визуальной аналитики составляют методы рисования и визуализации графов. Эта область является одной из наиболее наукоемких областей современных информационных технологий. Для работы в области визуализации информации, помимо знакомства с теорией графов, необходимо достаточно глубокое знакомство с компьютерной геометрией, компьютерной графикой и методами вычислений.

Алгоритмы визуализации графов имеют огромное количество областей применения. Большинство людей сталкивались с файловой иерархией на компьютере. Эта иерархия может быть изображена в виде дерева, которое является частным случаем графа. Другие знакомые типы графов включают иерархии, представимые организационными диаграммами и таксономии, которые отражают отношения между разными видами.

Наукоемкие продукты, использующие методы визуализации информации, широко распространяются на мировом рынке в течение последних 10-15 лет. Имеются фирмы, поставляющие библиотеки и программные комплексы, ориентированные на визуализацию графов общего назначения, а также системы визуализации программного обеспечения. Вместе с тем появляется все больше фирм, специализирующихся на визуализации бизнес-информации, необходимой аналитикам различных предприятий и ориентированной на профиль тех или иных предприятий.

Алгоритм для «чайников» стадий структурирования знаний (тут ниже график, так что можно не читать, а просто по графику запомнить)

В качестве простейшего прагматического подхода к формированию поля знаний начинающему инженеру по знаниям можно предложить следующий алгоритм для «чайников» (рис. 19).

1. Определение входных {X} и выходных {Y} данных. Этот шаг совершенно необходим, так как он определяет направление движения в поле знаний — от X к Y. Кроме того, структура входных и выходных данных существенно влияет на форму и содержание поля знаний. На этом шаге определение может быть достаточно размытым, в дельнейшем оно будет уточняться.

2. Составление словаря терминов и наборов ключевых слов N. На этом шаге проводится текстуальный анализ всех протоколов сеансов извлечения знаний, и выписываются все значимые слова, обозначающие понятия, явления, процессы, предметы, действия, признаки и т. п. При этом следует попытаться разобраться в значении терминов. Важен осмысленный словарь.

3. Выявление объектов и понятий {A}. Производится «просеивание» словаря N и выбор значимых для принятия решения понятий и их признаков. В идеале на этом шаге образуется полный систематический набор терминов из какой-либо области знаний.

4. Выявление связей между понятиями. Все в мире связано. Но определить, как направлены связи, что ближе, а что дальше, необходимо на этом этапе. Таким образом, стоится сеть ассоциаций, где связи только намечены, но пока не поименованы. Например, понятия «день», «ночь», «утро» и «вечер» явно как-то связаны, связаны также и понятия «красный флаг» и «красный галстук», но характер связи тут существенно отличен.

5. Выявление метапонятий и детализации понятий. Связи, полученные на предыдущем шаге, позволяют инженерам по знаниям структурировать понятия и как выявлять понятия более высокого уровня обобщения (метапонятия), так и детализировать на более низком уровне.

6. Построение пирамиды знаний. Под пирамидой знаний мы понимаем иерархическую лестницу понятий, подъем по которой означает углубление понимания и повышения уровня абстракции (обобщенности) понятий. Количество уровней в пирамиде зависит от особенностей предметной области, профессионализма экспертов и инженеров по знаниям.

7. Определение отношений {RA}. Отношения между понятиями выявляются как внутри каждого из уровней пирамиды, так и между уровнями. Фактически на это шаге даются имена тем связям, которые обнаруживаются на шагах 4 и 5, а также обозначаются причинно-следственные, лингвистические, временны и другие виды отношений.

8. Определение стратегий принятия решений (Sf). Определение стратегий принятия решения, то есть выявление цепочек рассуждений, связывает все сформированные ранее понятия и отношения в динамическую системы поля значений. Именно стратегии придают активность знаниям, именно они «перетряхивают» модель М в поиске от X к Y.



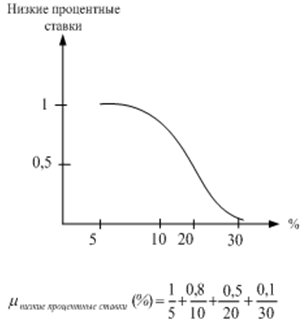
Рис. 19 - Стадии структурирования знаний — алгоритм для «чайников»

**25. Нечёткие и неопределённые знания. Нечёткие множества и лингвистические переменные. Вероятностные рассуждения**

Знания человека в большинстве случаев нечёткие. *Лингвистическая переменная* — в теории [нечётких множеств](https://ru.wikipedia.org/wiki/%D0%9D%D0%B5%D1%87%D1%91%D1%82%D0%BA%D0%BE%D0%B5_%D0%BC%D0%BD%D0%BE%D0%B6%D0%B5%D1%81%D1%82%D0%B2%D0%BE), [переменная](https://ru.wikipedia.org/wiki/%D0%9F%D0%B5%D1%80%D0%B5%D0%BC%D0%B5%D0%BD%D0%BD%D0%B0%D1%8F_%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D0%B0), которая может принимать значения [фраз](https://ru.wikipedia.org/wiki/%D0%A4%D1%80%D0%B0%D0%B7%D0%B0) из [естественного](https://ru.wikipedia.org/wiki/%D0%95%D1%81%D1%82%D0%B5%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) или [искусственного языка](https://ru.wikipedia.org/wiki/%D0%98%D1%81%D0%BA%D1%83%D1%81%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA). Высокая/средняя/низкая скорость, горячая/холодная/теплая/летняя вода.

Лотфи Заде в 1965 г. расширил классическое понятие множества, допустив, что характеристическая функция (функция *принадлежности элемента множеству*) может принимать любые значения в интервале [0,1], а не только значения 0 или 1. В основе данной теории лежит понятие функции принадлежности, которая указывает степень принадлежности какого-либо элемента некоторому множеству элементов. Данная функция является субъективной и строится на основании *знаний*, опыта или ощущений некоторого субъекта к какому-либо объекту, процессу, явлению и т.д.

, –Характ. функция, принимающая любое значение от 0 до 1. Множество М назы­вают множеством принадлежностей. Если М = {0, 1}, то нечеткое подмножество А может рассматриваться как обычное или четкое множество. Эти множества чаще всего представляют графиками, например

  
Основные положения нечеткой логики:

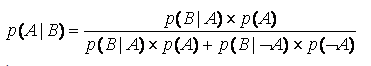
- истинность предпосылки, гипотезы или формулы лежит в интервале от 0 до 1;

- если две предпосылки (Е1 и Е2) соединены ∧ (логическим И), то истинность гипотезы Н рассчитывается по формуле t(Н) = MIN(t(Е1), t(Е2));

- если две предпосылки (Е1 и Е2) соединены ∨ (логическим ИЛИ), то истинность гипотезы Н рассчитывается по формуле t(Н) = MAX(t(Е1), t(Е2));

- если правило (П) имеет свою оценку истинности, тогда итоговая истинность гипотезы Нитог корректируется с учетом истинности правила t(Нитог) = MIN(t(Н), t(П))

Байесовскую теорию вероятности можно использовать в управлении неопределенностью. Оно обеспечивает путь для получения условной вероятности события B при условии А. Это соотношение позволяет управлять неопределенностью и «делать вывод «вперед и назад».



**26-27-28. Это можно вставить в три вопроса в начало!!! Дальше по подвопросам только инфа!!!**

*Иерархия многоуровневой модели приятия решений:*

1) уровень описания, или абстрагирования, - *страта*;

2) уровень сложности принимаемого решения - *слой*;

3) организационный уровень - *эшелон*.

Три введенных понятия *иерархических структур* имеют каждое свою область применения: *концепция страт* введена для целей моделирования, *концепция слоев* - для вертикальной декомпозиции решаемой проблемы на подпроблемы; *концепция эшелонов* относится к взаимной связи между образующими систему элементами принятия решения. Для больших систем эти концепции можно сливать. Например, одна страта может состоять из эшелонов.

**26. Основные виды иерархий. Уровень абстрагирования**

Страта. Для каждого уровня существует ряд характерных особенностей и переменных, законов и принципов, с помощью которых и описывается поведение системы. Чтобы такое описание было эффективным, необходима возможно большая независимость моделей для различных уровней системы. Для такой системы используется термин "*стратифицированная система*". Уровни абстрагирования будем называть "*стратами*". На каждой страте имеется свой собственный набор переменных, которые позволяют в значительной степени ограничиться изучением только одной *страты*. Но предположение о полной независимости страт было бы неоправданным и может привести к неполному пониманию системы в целом (нельзя ограничиться, например, исследованием человека только с биологической точки зрения).

Характеристики *стратифицированного описания систем*:

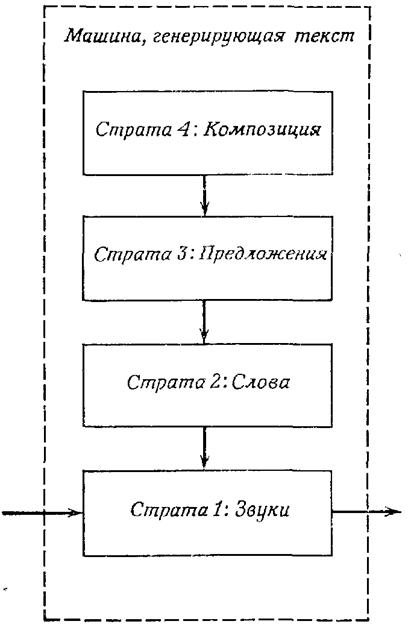
1. Выбор *страт* зависит от наблюдателя, его знания и заинтересованности в деятельности системы.

2. Принципы и законы, используемые для характеристики системы на любой *страте*, в общем случае не могут быть выведены из принципов, используемых на других стратах.

3. Требования, предъявляемые к работе системы на любой *страте*, выступают как условия или ограничения деятельности на нижестоящих *стратах*.

4. То, что является ответом рассмотрения на данной *страте*, более подробно раскрывается на нижерасположенной *страте*.

5. Чем ниже мы спускаемся по иерархии, тем более детальным становится раскрытие системы; чем выше мы поднимаемся, тем более явным становится смысл и значение всей системы.

  
Для правильного понимания сложной системы фундаментальную роль играет иерархический подход (стратифицированные модели). Вначале можно ограничиться, скажем, одной стратой, в зависимости от интересующей нас задачи и имеющегося запаса знаний, а затем можем либо детализировать свои знания, двигаясь вниз по иерархии, либо добиться более глубокого понимания системы, двигаясь вверх по иерархии. Выбор исходной страты отчасти определяется также простотой описания на ней.

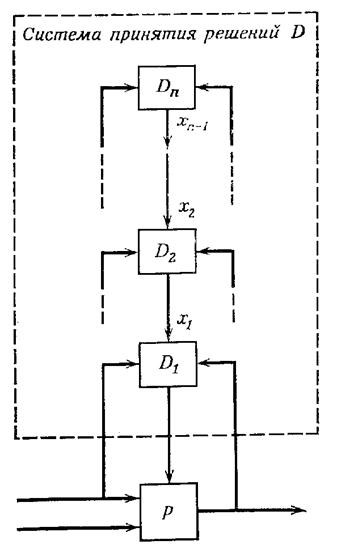
**27. Основные виды иерархий. Уровень сложности принимаемого решения**

Слой. Понятие слоя иерархии относится к *процессам принятия сложных решений*. В любой реальной ситуации принятия решения существуют две особенности:

1) когда приходит время принимать решение, принятие и выполнение решения желательно ускорить;

2) прежде чем принять решение, следует хорошо оценить создавшуюся ситуацию.

При принятии решения в сложных ситуациях разрешение этой дилеммы ищут в *иерархическом подходе*: определяют семейство проблем, которые пытаются разрешить последовательным путем в том смысле, что решение любой проблемы из этой последовательности определяет и фиксирует, какие-то параметры в следующей проблеме, так что последняя становится полностью определенной, и можно приступить к ее решению. Решение первоначальной проблемы достигнуто, как только решены все подпроблемы.

  
Каждый блок здесь представляет собой принимающий решение элемент. Выход элемента (например, D2) есть решение или последовательность решений задачи, зависящей от параметра, фиксируемого входом x2. Этот вход в свою очередь является выходом принимающего решение элемента более высокого уровня. Таким образом, сложная проблема принятия решения разбивается на семейство последовательно расположенных более простых подпроблем, так что решение всех подпроблем позволяет решить и исходную проблему. Такая иерархия–иерархией слоев принятия решений, а вся система принятия решений — многослойной системой принятия решений.

**28. Основные виды иерархий. Организационные иерархии (многоэшелонные системы)**

Эшелон. Это понятие иерархии подразумевает, что:

1) система состоит из семейства четко выделенных взаимодействующих подсистем;

2) некоторые из подсистем являются принимающими решения элементами;

3) принимающие решения элементы располагаются *иерархически*, т.е. некоторые из них находятся под влиянием или управляются другими решающими элементами.

Уровень в такой системе - *эшелон*.

Эти системы мы будем называть также многоэшелонными, многоуровневыми или многоцелевыми в связи с тем, что различные входящие в систему элементы, обладающие правом принятия решения, имеют обычно “конфликтные” цели. Это противоречие целей является не только побочным результатом эволюции и объединения различных подсистем в одну систему; можно показать, что оно даже необходимо для эффективного управления системой в целом.



По самой природе эшелонных многоуровневых многоцелевых систем элементы верхнего уровня в них хотя и обусловливают целенаправленную деятельность элементов нижних уровней, но не полностью управляют ею. Принимающим решения элементам нижних уровней должна быть предоставлена некоторая свобода в выборе их собственных решений; эти решения могут быть, но не обязательно будут, теми решениями, которые выбрал бы верхний уровень. Такая свобода действий — отличительная черта любой социальной или биологической многоуровневой системы. Можно показать, что для эффективного использования многоуровневой структуры существенно, чтобы элементам принятия решения была предоставлена некоторая свобода действий; должно быть проведено рациональное распределение усилий по принятию решений между элементами различных уровней.

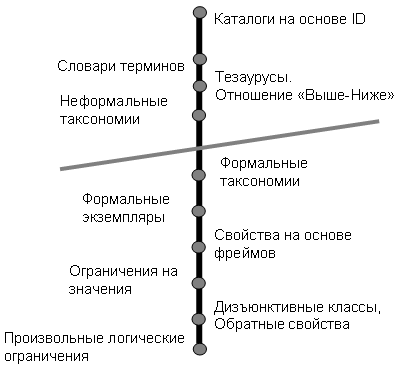
**29. Онтология для представления знаний. Классификация онтологий**

Онтология - это система, состоящая из набора понятий и набора утверждений об этих понятиях, на основе которых можно описывать классы, отношения, функции и индивиды.

Основными компонентами онтологии могут являться:

* Классы или понятия
* отношения или свойства, атрибуты
* Функции
* Аксиомы
* экземпляры или индивиды

Классификация онтологий по степени формальности.

Косая черта разделяет системы, предоставляющие "человеко-понятные" (выше черты) и "машино-понятные" (ниже черты) описания

***-****контролируемый словарь*, т.е. конечный список терминов (простейшим примером является каталог на основе идентификаторов).

***-****глоссарий*, представляющий собой список терминов с их значениями. Значения описываются в виде комментариев на естественном языке.

***-*** *Тезаурусы* несут дополнительную семантику, определяя связи между терминами. Отношения, свойственные для тезаурусов: синонимия, *иерархическое отношение* и ассоциация.

- *формальные таксономии*. Эта разновидность онтологий включает точное определение отношения ПОДКЛАСС-КЛАСС (обозначаемого как isA ). В таких системах строго соблюдается *транзитивность* отношения isA: если A является подклассом класса B, то каждый подкласс класса A также является подклассом класса B.

-формальные экзепляры. Наличие в онтологической системе формального отношения ЭКЗЕМПЛЯР-КЛАСС (обозначаемого как isInstanceOf ).

-*слоты*. Здесь классы (иногда их называют *фреймами* ) могут иметь информацию о *свойствах (слотах*).

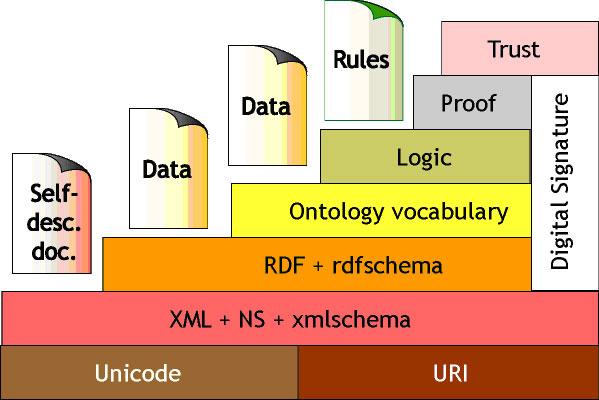
-Ограничения на значения и свойста(последние 3). Значения свойств берутся из некоторого предопределенного множества (целые числа, символьные константы) или из подмножества концептов онтологии (множество экземпляров данного класса, множество классов), вычленяются логически или методом исключения.

Классификация по цели создания:

* **Онтология представления.** Цель ее создания - описать область представления знаний, создать язык для спецификации других онтологий более низких уровней.
* Онтология верхнего уровня. Ее назначение - в создании единой "правильной" онтологии, фиксирующей знания, общие для нескольких предметных областей, и в многократном использовании данной онтологии.
* Онтология предметной области. Другое название - онтология домена. Назначение схоже с назначением онтологии верхнего уровня, но область интереса ограничена одной предметной областью (т.н. доменом), например, авиация, медицина, культура, дистанционное обучение, Интернет-технологии.
* Прикладная онтология. Назначение такой онтологий в том, чтобы описать концептуальную модель конкретной задачи или приложения.

**Классификация по содержимому**

* ***Общие онтологии*** описывают наиболее общие концепты (пространство, время, материя, объект, событие, действие и т.д.), которые независимы от конкретной проблемы или области. В эту категорию попадают и онтологии представления, и онтологии верхнего уровня.
* ***Онтология, ориентированная на задачу*** - это онтология, используемая конкретной прикладной программой и содержащая термины, которые используются при разработке ПО, выполняющего конкретную задачу. Она отражает специфику приложения, но может также содержать некоторые общие термины (например, в графическом редакторе будут и специфические термины - палитра, тип заливки, наложение слоев и т.д., и общие - сохранить и загрузить файл)
* ***Предметная онтология*** (или *онтология предметов*) описывает реальные предметы, участвующие в какой-либо деятельности (производстве). Например, это может быть онтология всех частей и компонентов самолетов определененной марки (Boeing) и сведения об их поставщиках, характеристиках, способе соединения друг с другом и т.п.

**30. Языки формализации онтологий**

Суть Semantic Web состоит в автоматизации "интеллектуальных" задач обработки значения тех ресурсов, имеющихся в Сети. Обработкой и обменом информации должны заниматься не люди, а специальные интеллектуальные агенты. Чтобы взаимодействовать между собой, агенты должны иметь общее формальное представление значения для любого ресурса. Именно для цели представления общей, явной и формальной спецификации значения в *Semantic* Web используются онтологии.

* RDF (*Resource Description Framework*). Язык описания ресурсов, основанный на триплетах "Субъект-Предикат-Объект" и спецификации URI. Концептуально RDF дает минимальный уровень для представления знаний в Сети. Спецификация RDF опирается на ранние стандарты, лежащие в основе Web, а именно Unicode, URI, XML и *XML Schema*.
* RDFS (*RDF Schema* ). Язык описания структурированных словарей для RDF. Он предоставляет минимальный набор средств для спецификации онтологий. Однако препятствием для *Semantic*Web стало то, что документов, написанных на языке RDF/RDFS, было относительно мало.
* GRDDL (Gleaning Resource Descriptions form *Dialects* of Languages). Его назначение состоит в предоставлении средств для извлечения RDF-триплетов из XML и XHTML данных (в особенности это относится к документам, автоматически генерируемым из закрытых баз данных).
* *OWL*. фактически является надстройкой над RDF/RDFS и поддерживает эффективное представление онтологий в терминах классов и свойств, обеспечение простых логических проверок целостности онтологии и связывание онтологий друг с другом (импорт внешних определений). Многие формализмы описания знаний могут быть отображены на формализм *OWL* (два из его диалектов - *OWL* *Lite* и *OWL* DL - соответствуют двум *дескриптивным логикам*, имеющим разную выразительную силу).
* *RIF* (Rule Interchange Format). Его назначение - соединить в одном стандарте несколько формализмов для описания правил (по которым может осуществляться нетривиальный логический вывод).
* Язык *SPARQL* - язык запросов к RDF-хранилищам. Синтаксически он очень похож на SQL.

31. История развития системы управления знаниями. Ресурсы знаний. Дуализм понятия «Управления знаниями»

Управление знаниями (knowledge management) — это концепция, которая появилась примерно два десятилетия назад, где-то в 1990х. определение термину **“управление знаниями**” — это организация и систематизация информации и знаний в компании. Томас Давенпорт предложил определение, которым пользуются до сих пор: “**Управление знаниями** — это процесс сбора, распространения и эффективного использования знаний”.

Знания вырабатываются, прежде всего, в сфере науки, точнее - в сфере научных исследований и опытно-конструкторских работ (НИОКР) и распространяются в основном через сферу образования и другие каналы, приобретая характер информации. Поэтому в современном мире велико значение научного потенциала (**научных ресурсов**), образовательного (**образовательных ресурсов**) и информационного (**информационных ресурсов**).

**Научные ресурсы** характеризуют объем и качество накопленных знаний и способность к их воспроизводству, их внедрению в виде инноваций. Масштабы НИОКР зависят от ресурсного обеспечения (финансирование, кадры, научное оборудование), организационной структуры, государственной политики.

**Важнейшими показателями, характеризующими научные ресурсы:**

* доля расходов на НИОКР в ВВП;
* численность занятых в НИОКР (исследователей и технического персонала) в общей численности населения;
* индекс цитирования (частота ссылок в научных изданиях на работы исследователей данной страны) с учетом языковых барьеров
* число международных премий (прежде всего Нобелевских) за выдающиеся научные достижения;
* сальдо технологического баланса, т.е. баланс торговли лицензиями и патентами;
* доля высокотехнологичных отраслей в промышленном экспорте

**Образовательные ресурсы** – накопленные поколениями объем и качество знаний и профессионального опыта, которые усвоены населением и воспроизводятся через систему образования.

Основные тенденции развития сферы образования таковы: построение непрерывной цепочки этапов образования, фундаментализация и профессионализация (сочетание базовых знаний со специальными), ускоренное развитие новых дисциплин (информатика, биотехнология и т.д.).

**Важнейшими показателями, характеризующими образовательные ресурсы отдельных стран, являются:**

* доля расходов на образование в ВВП;
* расходы на образование на душу населения;
* доля молодых людей, окончивших полную среднюю школу;
* доля студентов в возрастной группе 18-29 лет;
* численность студентов на 10 тыс. населения;
* доля лиц с высшим образованием;
* доля иностранцев в общей численности студентов.

Совокупность информации, предназначенной для распространения, приобретения и использования, находящейся как в государственной, так и негосударственных формах собственности формирует **национальные информационные ресурсы.**

**Важнейшими показателями информационных ресурсов отдельных** стран является доступ к информационными источникам как к традиционным (радио, телевидение, печать, стационарная телефония), так и новым (мобильная связь, Интернет ресурс):

* число ежедневных печатных изданий, радиоприемников, телевизоров, подключений к кабельному телевидению, число телефонных номеров в расчете на 100 % населения;
* число мобильных телефонов, серверов и пользователей Интернета в расчете на 100% населения;
* доля расходов комплекса информационно-коммуникационных отраслей (ИКТ) в ВВП;
* число персональных компьютеров в расчете на 100% населения;
* индекс «цифрового доступа» (характеризует возможности жителя страны получить доступ и использовать возможности ИКТ).



Фактически *«Управление знаниями»*  можно рассматривать и как новое направление в менеджменте, и как направление в информатике для поддержки процессов создания, распространения, обработки и использования знаний внутри предприятия

32. Жизненный цикл знаний в системах управления знаниями (СУЗ)

В соответствии с методологией системного анализа, в жизненном цикле знаний (ЖЦЗ), как и в жизненном цикле любой сферы деятельности, можно выделить три основных процесса (этапа):

* выявление потребности в знаниях;
* производство (создание) знаний;
* потребление (утилизация) знаний.

Так же жизненный цикл можно разделить на такие этапы:

* Определение потребности в знаниях.
* Определение источников знаний (определяете источники знаний по каждому виду знаний. Например, в качестве источников вы выбираете: 1) подразделение вашей компании, в которой добились выдающихся успехов в этом; 2) внешние журналы, сайты. Обзоры и исследования 3) группу специалистов компании, перед которыми будет поставлена задача адаптировать существующие знания к реальности вашей компании, а также в случае необходимости создавать еще несуществующие знания.)
* создание нового знания.
* Формализация знаний (Все создаваемые знания вы формализуете в виде Технических условий, описания технологических процессов)
* Сохранение и структуризация знаний
* Распространение и использование
* Коррекция содержания и практики использования знаний (Оценка практики использования знаний дополняется мониторингом результатов деятельности каждого участка. Случаи отклонения от нормативов рассматриваются как с точки зрения выяснения способов повышения эффективности производственного процесса и распространения лучшего опыта на другие участки, так и с точки зрения обучения и развития отстающих подразделений.)

В целом процесс создания СУЗ подразумевает наличие постоянно-действующего подразделения или рабочей группы, поддерживающей систему на «плаву».

33. Система управления знаниями. Корпоративная память 

Для управления знаниями используются следующие информационные технологии:

* • электронная почта (e-mail);
* • базы данных и хранилища данных (data base & data warehouse);
* • системы групповой поддержки (groupware);
* • браузеры и системы поиска (browsers);
* • корпоративные сети и Интернет (Intranet & Internet);
* • экспертные системы и базы знаний (expert systems & knowledge base);
* • интеллектуальные информационные системы (intelligence systems);
* • системы искусственного интеллекта (artificial intelligence systems).

Из чего состоит система управления знаниями? Очевидный ответ: из информации и данных, которые доступны всем членам организации через специальные порталы и системы управления контентом.

**Отличительной особенностью системы управления** знаниями является интеграция множества разнородных, часто территориально распределенных источников знаний для решения общих задач. СУЗ интегрирует знания как из внутренних, так и из внешних источников. Источники знаний могут иметь недокументированную форму (неявные знания экспертов), документированную текстовую, табличную, графическую и структурируемую форму в виде баз знаний экспертных систем.

Система управления знаниями обычно используется в двух аспектах:

— обеспечение качественными знаниями процессов решения различных задач;

— создание интерактивной среды взаимодействия специалистов в процессе решения задач.

**Интеллектуальная информационная система (intelligent system), иногда называемая системой, основанной на знаниях**, представляет собой комплекс программных, лингвистических и логико-математических средств для реализации основной задачи – осуществления поддержки деятельности человека. В число задач, решаемых интеллектуальной информационной системой, входят: интерпретация данных, диагностика, мониторинг, проектирование, прогнозирование, планирование, обучение, управление, поддержка принятия решений.

Одним из решений по управлению знаниями является понятие **корпоративной памяти (corporate memory)**, которая по аналогии с человеческой памятью позволяет пользоваться предыдущим опытом и избегать повторения ошибок.

Корпоративная память фиксирует информацию из различных источников предприятия и делает эту информацию доступной специалистам для решения производственных задач. Корпоративная память не позволяет исчезнуть знаниям выбывающих специалистов (уход на пенсию, увольнение и пр.). Она хранит большие объемы данных, информации и знаний из различных источников предприятия. Они представлены в различных формах, таких как базы данных, архивы документов и базы знаний.

**Разработка систем корпоративной памяти включает в себя пять этапов:**

1) накопление – стихийное и бессистемное накопление информации в организации;

2) извлечение – процесс, идентичный традиционному извлечению знаний для экспертных систем.

3) структурирование – выделение основных понятий, выработка структуры представления информации, максимальная наглядность, простота изменения и дополнения;

4) формализация – представление структурированной информации в форматах машинной обработки, т.е. на языках описания данных и знаний;

5) обслуживание – корректировка формализованных данных и знаний (добавление, обновление): удаление устаревшей информации; фильтрация данных и знаний для поиска.

Автоматизированные системы корпоративной памяти (organizational memory information systems – OMIS) предназначены для накопления и управления знаниями предприятия. Эти системы включают в себя работу как на уровне с явным знанием компании в форме баз данных и электронных архивов, так и на уровне со скрытым знанием, фиксируя его в некотором (более или менее формальном) представлении в форме экспертных систем или баз данных. Автоматизированные системы корпоративной памяти часто используют вспомогательные справочные системы, так называемые helpdesk-приложения.

**Основными функциями OMIS являются:**

• сбор и систематическая организация информации из различных источников в централизованное и структурное информационное хранилище;

• интеграция с существующими автоматизированными системами. На техническом уровне это означает, что корпоративная память должна быть непосредственно связана с помощью интерфейса с инструментальными средствами, которые в настоящее время используются в организации (например, текстовые процессоры, электронные таблицы, системы);

• обеспечение нужной информации по запросу (пассивная форма) и при необходимости (активная форма). Слишком частые ошибки – следствие недостаточной информированности. Этого невозможно избежать с помощью пассивной информационной системы, так как служащие часто слишком заняты, чтобы искать информацию, или просто не знают, что нужная информация существует. Корпоративная память может напоминать служащим о полезной информации и быть компетентным партнером для совместного решения задач.

34. Системы Business Intelligence. История появления Business Intelligence. Основные понятия и особенности 

Впервые термин «business intelligence» был введен в обращение аналитиками Gartner в конце 1980-х годов, как «пользователецентрический процесс, который включает доступ и исследование информации, ее анализ, выработку интуиции и понимания, которые ведут к улучшенному и неформальному принятию решений». Позже в 1996 году появилось уточнение — «инструменты для анализа данных, построения отчетов и запросов могут помочь бизнес-пользователям преодолеть море данных, — сегодня эти инструменты в совокупности попадают в категорию, называемую бизнес-интеллект (Business Intelligence)».

**Business Intelligence** (бизнес-аналитика, бизнес-анализ) - программное обеспечение, созданное для помощи управленцу в анализе информации о своей компании и её окружении. BI-технологии позволяют анализировать большие объёмы информации, заостряя внимание пользователей лишь на ключевых факторах эффективности, моделируя исход различных вариантов действий, отслеживая результаты принятия тех или иных решений.

Сегодня под Business Intelligence **понимают** главным образом технологии, связанные с хранением и анализом фактографической структурированной информации (базы данных, плоские файлы и т.п.) и квазиструктурированной информации (XML).

Business Intelligence имеет отношение к процессу превращения данных в знания, а знаний в действия бизнеса для получения выгоды. Является деятельностью конечного пользователя, которую облегчают различные аналитические и групповые инструменты и приложения, а также инфраструктура хранилища данных.

**Инструменты Business Intelligence** — программное обеспечение, которое позволяет бизнес-пользователям видеть и использовать большое количество сложных данных.

Системы BI:

**Cognos от IBM**. Новое программное обеспечение поддерживает естественную сквозную цепочку информационных представлений, от просмотра и изучения до принятия решения.

**Microsoft BI.** Удачная ценовая политика и интеграция с MS Office делает решения Microsoft особенно привлекательными для организаций, которые базируются на инфраструктурных решениях этой компании.

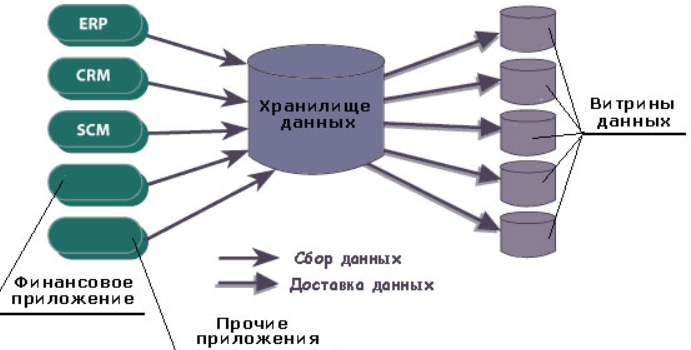
**Oracle** . Сильные стороны Essbase OLAP и возможности интеграции Hyperion с Microsoft Office повышают потенциал Oracle на рынке BI.

**ОСОБЕННОСТИ**

* интегрированная платформа. BI представляет собой комплексную платформу для управления интеллектуальными ресурсами предприятия со встроенными аналитическими инструментами, в числе которых оперативная аналитическая обработка данных (OLAP), средства внесения и обновления данных, средства извлечения, преобразования и загрузки данных, хранилища данных и функции генерации отчетов. Этот комплексный, интегрированный подход позволяет организациям легко создавать и развертывать мощные системы для управления интеллектуальными ресурсами предприятия, контролируя свои затраты.
* безопасность и эксплуатационная готовность. Параметры масштабируемости, эксплуатационной готовности и безопасности помогают обеспечить пользователям бесперебойный доступ к системам
* возможности анализа в масштабе всего предприятия. Усовершенствованное средство извлечения, преобразования и загрузки данных облегчает организациям объединение и анализ данных из множества разнородных источников.
* кардинальное сокращение затрат на сбор и обработку информации в компаниях с распределенной структурой
* повышение оперативности и качества подготовки управленческих решений
* автоматизированный контроль основных показателей деятельности предприятия.
* оперативное информирование заинтересованных лиц об отклонениях ключевых показателей от заданных нормативов
* оптимальное использование вычислительных ресурсов компании, например, путем снижения нагрузки на расчётные системы
* конфиденциальность информации, разграничение и регистрация доступа к данным.
* **Дополнительные достоинства**: гибкая адаптация к потребностям пользователей, удобный интерфейс, презентабельные отчёты, богатый аналитический инструментарий, удалённый доступ через сеть Интернет, относительно быстрое внедрение, техническая поддержка на этапе эксплуатации и возможности дальнейшего совершенствования.

35. Системы Business Intelligence. Архитектура и жизненный цикл 

**Архитектура.** Данные стекаются из разных систем в Хранилище данных и потом идут в витрины данных.

*  ETL-инструменты: программы, позволяющие выполнять загрузку данных в DWH из различных учетных систем.
* DWH-хранилище: полноценная база данных SQL для подготовки и хранения данных для аналитики.
* OLAP-кубы: технология, позволяющая делать в реальном времени (1-5 секунд) любые отчеты и проводить полноценный анализ данных.
* Клиентские приложения: как правило, для детального анализа данных и построения динамических отчетов пользователи используют Сводные таблицы Microsoft Excel, подключенные к OLAP-кубам. Для поверхностного анализа и визуализации ключевых показателей также используются WEB-приложения, которые должны поддерживать доступ к отчетам с любого устройства: компьютер, планшет, телефон.

**Жизненный цикл:**

1) сбор и первичная обработка данных; (К первому уровню архитектуры относятся упоминавшиеся уже источники данных, как правило именуемые транзакционными или операционными источниками (базами) данных, являющимися частью так называемых **OLTP**-систем (online transactional processing).)

2) извлечение, преобразование и загрузка данных;( Процесс извлечения, преобразования и загрузки данных поддерживается так называемыми **ETL**-инструментами (extraction, transformation, loading), предназначенными для извлечения данных из различных транзакционных источников нижнего уровня, их преобразования и консолидации, а также загрузки в целевые аналитические базы данных - хранилища данных и витрины данных.)

3) складирование данных; (К третьему уровню архитектуры относятся источники данных, которые называют хранилищами данных (от англ. **Data Warehouse**). Хранилища данных включают в себя источники данных, ориентированные на хранение и анализ информации. Такие источники могут объединять информацию из нескольких транзакционных систем и позволяют анализировать ее в комплексе с применением современных программных инструментов делового анализа данных.)

4) представление данных в витринах данных;(К четвертому уровню архитектуры ИАС относятся источники данных, называемые **витринами данных (data marts**), предназначенные для проведения целевого делового анализа. Многомерные витрины организуются в виде многомерных баз данных **OLAP** (Online Analytical Processing), где справочная информация представляется в виде измерений, а количественная - в виде показателей.)

5) анализ данных; (К следующему уровню архитектуры организации относятся современные программные средства, именуемые инструментами интеллектуального или делового анализа данных (Business Intelligence Tools), или BI-инструменты.Инструменты интеллектуального анализа данных используются пользователями для доступа к информации, ее визуализации, многомерного анализа и формирования как предопределенных по форме и составу, так и произвольных отчетов)

6) Web-портал.(доступа к информации через привычный Web-браузер позволяет экономить на затратах, связанных с закупкой и поддержкой настольных аналитических приложений для большого числа клиентских мест. Реализация Web-портала позволяет снабжать аналитической информацией как пользователей внутри офиса, так и мобильных пользователей-аналитиков)

36. Обработка естественного языка в интеллектуальных системах

**Основной проблемой обработки ПМ является языковая неоднозначность.**

**Обработка естественного языка** (Natural Language Processing, NLP) — общее направление искусственного интеллекта и математической лингвистики. Оно изучает проблемы компьютерного анализа и синтеза естественных языков.

**задачи обработки естественного языка. Это:**

1) распознавание речи;( преобразованию речевого сигнала человеческого голоса в цифровую информацию)

2) синтез речи;(формирование речевого сигнала по печатному тексту)

3) анализ текста;( извлечения содержательной, высокого качества информации из текста на естественном языке)

4) синтез текста;

5) машинный перевод;(процесс перевода устных текстов, написанных на естественном языке, на другой, тоже естественный, язык при помощи электронно-вычислительных машин)

6) создание вопросно-ответных систем;(принимать, распознавать, классифицировать вопросы и давать ответы на них на естественном языке.1.определения типа вопроса 2.поиском текстов, потенциально содержащих ответ на этот вопрос;3.извлечением ответа из этих источников.)

7) информационный поиск;( 1.формулирование информационного запроса; 2.поиск потенциальных обладателей соответствующей информации;3.извлечение информации из найденных документов;4.ознакомление с результатами произведённого поиска и выборка наиболее подходящих условиям запроса источниками.)

8) извлечение информации;

9) анализ тональности текста(анализ лексем текста, оценка их эмоциональной окрашенности и классификация по принадлежности к нейтральному, позитивному или негативному лексическому слою языка)

10) реферирование текста.(сокращение объёма текста за счёт выделения основных тезисов путём поиска соответствий заданным в поиске ключевым словам и его краткое изложение)

пакеты программного обеспечения, способного выполнить тот или иной тип обработки естественного языка:

• General Architecture for Text Engineering (GATE)(когда требуется определить смысловое содержание некого текста и кодировать его путем добавления аннотаций к отдельным сегментам текста в структурированном виде. Эта система применяется для выполнения задач по извлечению информации,)

• ModularAudioRecognitionFramework(MARF)(модульная библиотека распоз навания аудио. Данное программное обеспечение предназначено для выборки алгоритмов для обработки и распознавания текста, речи, звука, то есть выполнения непосредственных потребностей поставленных задач обработки естественного языка)

• NaturalLanguageToolkit (NLTK).(библиотеки и программ для символьной и статистической обработки естественного языка.)