



МОДЕЛИ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ

- ❖ Извлечение информации.
- ❖ Обработка информации.
- ❖ Хранение информации.
- ❖ Представление и использование информации.

Представления предметной области

Любая предметная область рассматривается в виде трех представлений:

- **реальное** представление предметной области
- **формальное** представление предметной области
- **информационное** представление предметной области



Извлечение информации

При извлечении информации важное место занимают различные формы и методы исследования данных:

- поиск ассоциаций, связанных с привязкой к какому-либо событию;
- обнаружение последовательностей событий во времени;
- выявление скрытых закономерностей по наборам данных, путем определения причинно-следственных связей между значениями определенных косвенных параметров исследуемого объекта (ситуации, процесса);
- оценка важности (влияния) параметров на развитие ситуации;
- классифицирование (распознавание), осуществляемое путем поиска критериев, по которым можно было бы относить объект (события, ситуации, процессы) к той или иной категории;
- кластеризация, основанная на группировании объектов по каким-либо признакам;
- прогнозирование событий и ситуаций.



Извлечение информации

ДО

ФИО	Роль	Организация

Президиум Российской академии наук решил назначить члена-корреспондента РАН, директора Курчатовского института Михаила Ковальчука исполняющим обязанности вице-президента РАН.

Как пояснил секретарь общественной палаты Евгений Велихов, решением президиума Ковальчук назначается лишь исполняющим обязанности, а потом он "будет избираться на этот пост".

Кандидатуру Михаила Ковальчука на пост вице-президента РАН предложил лично президент академии Юрий Осипов, пишет газета "Коммерсант".

ПОСЛЕ

ФИО	Роль	Организация
Михаил Ковальчук	директор	Курчатовский институт
Евгений Велихов	секретарь	общественная палата
Юрий Осипов	президент	академия

Президиум Российской академии наук решил назначить члена-корреспондента РАН, **директора Курчатовского института Михаила Ковальчука** исполняющим обязанности вице-президента РАН.

Как пояснил **секретарь общественной палаты Евгений Велихов**, решением президиума Ковальчук назначается лишь исполняющим обязанности, а потом он "будет избираться на этот пост".

Кандидатуру Михаила Ковальчука на пост вице-президента РАН предложил лично **президент академии Юрий Осипов**, пишет газета "Коммерсант".

Декомпозиция

Декомпозиция – научный метод, использующий структуру задачи и позволяющий заменить решение одной большой задачи решением серии меньших задач.



Основные понятия декомпозиции

Декомпозиция на основе объектно-ориентированного подхода основана на выделении следующих основных понятий:

- ❖ Объект
- ❖ Класс
- ❖ Экземпляр



- **Объект** – это абстракция множества предметов реального мира, обладающих одинаковыми характеристиками и законами поведения. Объект характеризует собой типичный неопределенный элемент такого множества.
- Основной **характеристикой** объекта является состав его атрибутов (свойств).
- **Атрибуты** – это специальные объекты, посредством которых можно задать правила описания свойств других объектов.

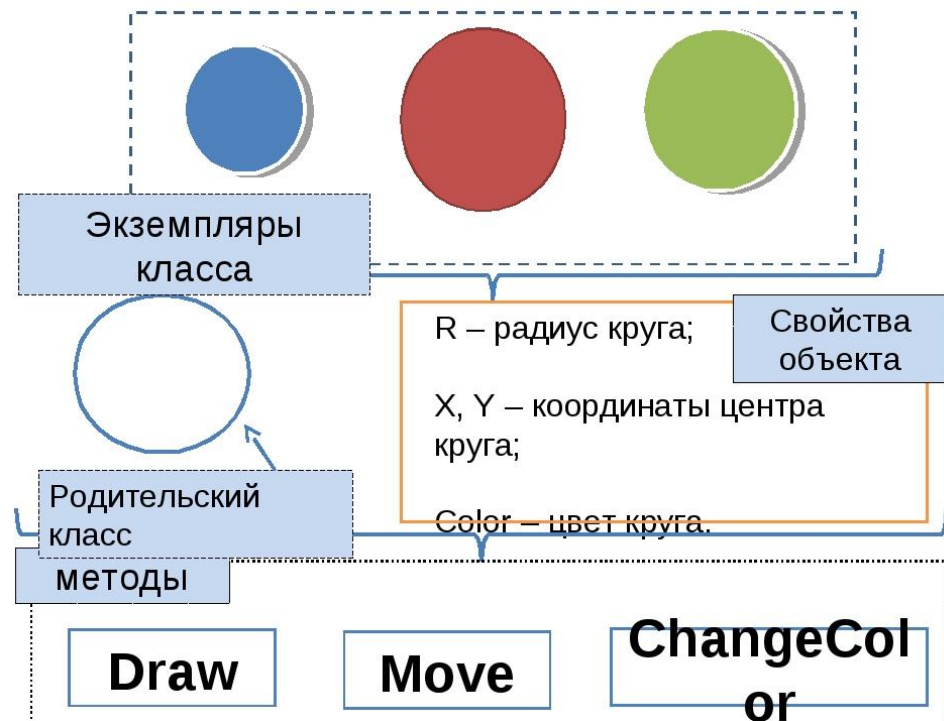


Основные понятия декомпозиции

- **Экземпляр** объекта – это конкретный элемент множества.

Например, объектом может являться государственный номер автомобиля, а экземпляром этого объекта – конкретный номер.

- **Класс** – это множество предметов реального мира, связанных общностью структуры и поведением. Элемент класса – это конкретный элемент данного множества.



Эволюция объектно-ориентированной системы

- - подразумевает определение новых классов объектов на основе базовых

- ❖ **Полиморфизм** – способность объекта принадлежать более чем одному типу.
- ❖ **Наследование** выражает возможность определения новых классов на основе существующих с возможностью добавления или переопределения данных и методов.
- ❖ **Инкапсуляция** – это процесс отделения друг от друга элементов объекта, определяющих его устройство и поведение; инкапсуляция служит для того, чтобы изолировать контрактные обязательства абстракции от их реализации.

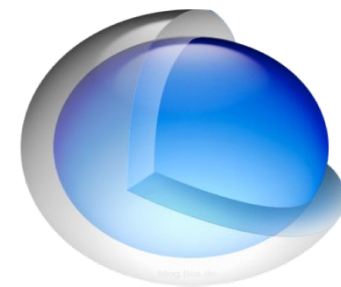
Методы обогащения информации



- ❖ Структурное
 - ❖ Статистическое
 - ❖ Семантическое
 - ❖ Прагматическое
- **Структурное обогащение** предполагает изменение параметров сообщения, отображающего информацию в зависимости от частотного спектра исследуемого процесса, скорости обслуживания источников информации и требуемой точности.
 - При **статистическом обогащении** осуществляют накопление статистических данных и обработку выборок из генеральных совокупностей накопленных данных.

Методы обогащения информации

- **Семантическое обогащение** означает минимизацию логической формы, исчислений и высказываний, выделение и классификацию понятий, содержания информации, переход от частных понятий к более общим. В итоге семантического обогащения удастся обобщенно представить обрабатываемую либо передаваемую информацию и устранить логическую противоречивость в ней.
- **Прагматическое обогащение** является важной ступенью при использовании информации для принятия решения, при котором из полученной информации отбирается наиболее ценная, отвечающая целям и задачам пользователя.



Обработка информации

- **Обработка информации** состоит в получении одних «информационных объектов» из других «информационных объектов» путем выполнения некоторых алгоритмов и является одной из основных операций, осуществляемых над информацией, и главным средством увеличения ее объема и разнообразия.

Обработка информации



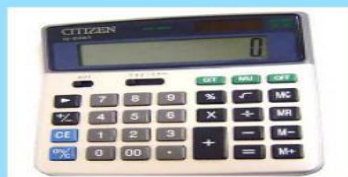
Русские счёты



Бухгалтерские счёты
(1970-е гг.)



Арифмометр Однера



Виды обработки информации

С точки зрения реализации на основе современных достижений вычислительной техники выделяют следующие виды обработки информации:

- ❖ **последовательная обработка**, применяемая в традиционной фоннеймановской архитектуре ЭВМ, располагающей одним процессором;
- ❖ **параллельная обработка**, применяемая при наличии нескольких процессоров в ЭВМ;
- ❖ **конвейерная обработка**, связанная с использованием в архитектуре ЭВМ одних и тех же ресурсов для решения разных задач, причем если эти задачи тождественны, то это **последовательный конвейер**, если задачи одинаковые – **векторный конвейер**.



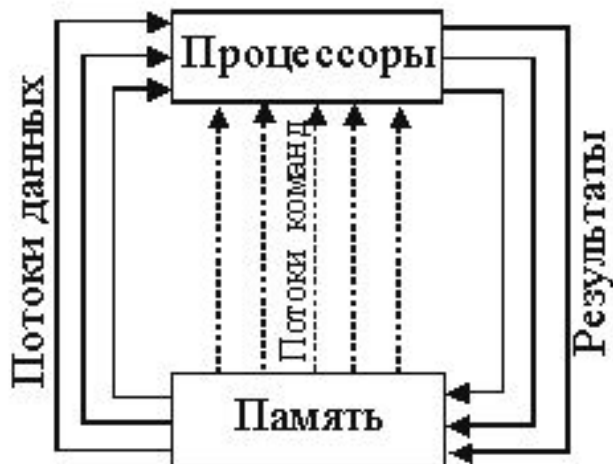
Классификация архитектур ЭВМ

- **Архитектуры с одиночным потоком команд и данных (SISD).** Традиционная архитектура фон Неймана + КЭШ + память + конвейеризация
- **Архитектуры с одиночными потоками команд и данных (SIMD).** Особенностью данного класса является наличие одного (центрального) контроллера, управляющего рядом одинаковых процессоров.



Классификация архитектур ЭВМ по способу обработки информации

- Архитектуры с множественным потоком команд и одиночным потоком данных (MISD). Один из немногих – систолический массив процессоров, в котором процессоры находятся в узлах регулярной решетки, роль ребер которой играют межпроцессорные соединения.



Архитектуры с множественным потоком команд и множественным потоком данных (MIMD). К этому классу могут быть отнесены следующие конфигурации: мультипроцессорные системы, системы с мультобработкой, вычислительные системы из многих машин, вычислительные сети.

Условия протекания процесса принятия решения

В зависимости от степени информированности о состоянии управляемого процесса, полноты и точности моделей объекта и системы управления, взаимодействия с окружающей средой, процесс принятия решения протекает в различных условиях:

1. Принятие решений в условиях определенности.
2. Принятие решений в условиях риска.
3. Принятие решений в условиях неопределенности.
4. Принятие решений в условиях многокритериальности.



Принятие решений в условиях определенности

- В этой задаче модели объекта и системы управления считаются заданными, а влияние внешней среды – несущественным.



Поэтому между выбранной стратегией использования ресурсов и конечным результатом существует однозначная связь, откуда следует, что в условиях определенности достаточно использовать решающее правило для оценки полезности вариантов решений, принимая в качестве оптимального то, которое приводит к наибольшему эффекту.

Принятие решений в условиях риска

- В отличие от предыдущего случая для принятия решений в условиях риска необходимо учитывать влияние внешней среды, которое не поддается точному прогнозу, а известно только вероятностное распределение ее состояний.



В этих условиях использование одной и той же стратегии может привести к различным исходам, вероятности появления которых считаются заданными или могут быть определены.

Принятие решений в условиях неопределенности

- Между выбором стратегии и конечным результатом отсутствует однозначная связь.



Кроме того, неизвестны также значения вероятностей появления конечных результатов, которые либо не могут быть определены, либо не имеют в контексте содержательного смысла.

Принятие решений в условиях многокритериальности

- В любой из перечисленных выше задач многокритериальность возникает в случае наличия нескольких самостоятельных, не сводимых одна к другой целей.



Наличие большого числа решений усложняет оценку и выбор оптимальной стратегии. Одним из возможных путей решения является использование методов моделирования.

Для решения задач в экспертных системах используют:

- ❖ метод логического вывода, основанный на технике доказательств, называемой резолюцией и использующей опровержение отрицания (доказательство «от противного»);
- ❖ метод структурной индукции, основанный на построении дерева принятия решений для определения объектов из большого числа данных на входе;
- ❖ метод эвристических правил, основанных на использовании опыта экспертов, а не на абстрактных правилах формальной логики;
- ❖ • метод машинной аналогии, основанный на представлении информации о сравниваемых объектах в удобном виде, например, в виде структур данных, называемых фреймами.



Условия принятия решений

Компоненты поддержки принятия решений:

- обобщающий анализ
- прогнозирование
- ситуационное моделирование

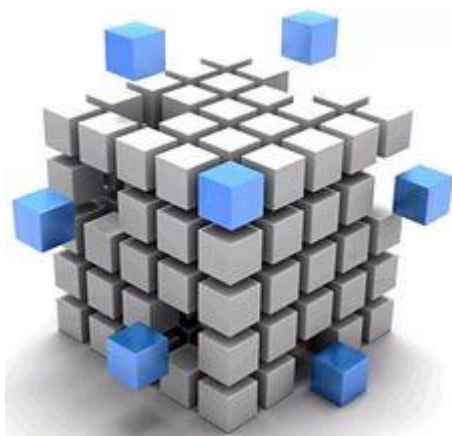


Аналитические системы поддержки принятия решений (СППР) позволяют решать три основных задачи:

- ❖ ведение отчётности
- ❖ анализ информации в реальном времени (OLAP)
- ❖ интеллектуальный анализ данных.

Системы поддержки принятия решений

OLAP (On-Line Analytical Processing) – сервис представляет собой инструмент для анализа больших объемов данных в режиме реального времени.



- ❑ Взаимодействуя с OLAP-системой, пользователь сможет осуществлять гибкий просмотр информации, получать произвольные срезы данных, и выполнять аналитические операции детализации, свертки, сквозного распределения, сравнения во времени.
- 】 В зависимости от функционального наполнения интерфейса системы выделяют два основных типа систем поддержки принятия решений: EIS и DSS.

Системы поддержки принятия решений

- ❖ **EIS (Execution Information System)** – информационные системы руководства предприятия. Эти системы ориентированы на неподготовленных пользователей, имеют упрощенный интерфейс, базовый набор предлагаемых возможностей, фиксированные формы представления информации.
- ❖ **DSS (Decision Support System)** – полнофункциональные системы анализа и исследования данных, рассчитанные на подготовленных пользователей, имеющих знания как в части предметной области исследования, так и в части компьютерной грамотности.



Системы поддержки принятия решений

- ❖ **PS (Presentation Services)** – средства представления. Обеспечиваются устройствами, принимающими ввод от пользователя и отображающими то, что сообщает ему компонент логики представления PL, плюс соответствующая программная поддержка.
- ❖ **PL (Presentation Logic)** – логика представления. Управляет взаимодействием между пользователем и ЭВМ. Обрабатывает действия пользователя по выбору альтернативы меню, по нажатию кнопки или выбору элемента из списка.
- ❖ **BL (Business or Application Logic)** – прикладная логика. Набор правил для принятия решений, вычислений и операций, которые должно выполнить приложение.



Системы поддержки принятия решений

- ❖ **DL (Data Logic)** – логика управления данными. Операции с базой данных (SQL-операторы SELECT, UPDATE и INSERT), которые нужно выполнить для реализации прикладной логики управления данными.
- ❖ **DS (Data Services)** – операции с базой данных. Действия СУБД, вызываемые для выполнения логики управления данными, такие как манипулирование данными, определения данных, фиксация или откат транзакций и т.п. СУБД обычно компилирует SQL-приложения.
- ❖ **FS (File Services)** – файловые операции. Дисковые операции чтения и записи данных для СУБД и других компонент. Обычно являются функциями ОС.

DATA LOGIC DL
Bar Code & More



Хранение информации

- **Хранение и накопление** являются одними из основных действий, осуществляемых над информацией и главным средством обеспечения ее доступности в течение некоторого промежутка времени.



- **База данных** может быть определена как совокупность взаимосвязанных данных, используемых несколькими пользователями и хранящихся с регулируемой избыточностью.

Компоненты базы данных

- **Банк данных** – система, представляющая определенные услуги по хранению и поиску данных определенной группе пользователей по определенной тематике.
- **Система баз данных** – совокупность управляющей системы, прикладного программного обеспечения, базы данных, операционной системы и технических средств, обеспечивающих информационное обслуживание пользователей.
- **Хранилище данных** (ХД – используют также термины Data Warehouse, «склад данных», «информационное хранилище») – это база, хранящая данные, агрегированные по многим измерениям.
- **Витрины данных** – множество тематических БД, содержащих информацию, относящуюся к отдельным информационным аспектам предметной области.

Программная реализация СУБД

Программная реализация (в дальнейшем СУБД) представляет собой набор программных модулей, работает под управлением конкретной ОС и выполняет следующие функции:

- описание данных на концептуальном и логическом уровнях;
- загрузку данных;
- хранение данных;
- поиск и ответ на запрос (транзакцию);
- внесение изменений;
- обеспечение безопасности и целостности.



языковые средства:

- ❖ язык описания данных (ЯОД);
- ❖ язык манипулирования данными (ЯМД);
- ❖ прикладной (встроенный) язык данных (ПЯД, ВЯД).

ПРЕДСТАВЛЕНИЕ И ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИИ



Описание предметной области

- ❑ **Концептуальный уровень** связан с частным представлением данных группы пользователей в виде внешней схемы, объединяемых общностью используемой информации.
- ❑ **Логический уровень** является обобщенным представлением данных всех пользователей в абстрактной форме.



Используются три вида моделей:

- Иерархические
- Сетевые
- Реляционные

- ❑ **Физический (внутренний) уровень** связан со способом фактического хранения данных в физической памяти ЭВМ. Основными компонентами физического уровня являются хранимые записи, объединяемые в блоки; указатели, необходимые для поиска данных; данные переполнения; промежутки между блоками; служебная информация.

Классификация БД

по способу хранения информации:

- интегрированные;
- распределенные;

по типу пользователя:

- монопользовательские;
- многопользовательские;



по характеру использования данных:

- монопользовательские;
- предметные.

Задачи представления информации

- Основной задачей операции представления информации пользователю является создание эффективного интерфейса в системе «человек—компьютер». При этом осуществляется преобразование информации в форму, удобную для восприятия пользователя.



Среди существующих вариантов интерфейса в системе «человек—компьютер» можно выделить два основных типа:

- ❖ на основе меню («смотри и выбирай»)
- ❖ на основе языка команд («вспоминай и набирай»).

Представление информации

Технология представления информации должна давать дополнительные возможности для понимания данных пользователями, поэтому целесообразно использование графики, таблиц, диаграмм, карт и т.д.



В сценарии работы пользователя при информационном наполнении понятий предметной области выделяем две фазы:

- ❖ выбор окон;
- ❖ работа с окнами.



Функции фаз выбора объектов:

- работа с общим каталогом окон в главном разделе
- создание нового раздела
- удаление раздела
- редактирование описания раздела
- передача определений и окон между разделами
- движение по иерархии разделов
- отбор разделов для работы
- отбор окон для работы





Спасибо за внимание