

Глава 4

Базовые информационные процессы, их характеристика и модели.

Информационные технологии основаны на реализации информационных процессов, разнообразие которых требует выделения базовых. К ним можно отнести извлечение, транспортирование, обработку, хранение, представление и использование информации. На логическом уровне должны быть построены математические модели, обеспечивающие параметрическую и критериальную совместимость информационных процессов в системе информационных технологий.

В процессе извлечения информации основной акцент сделан на формы и методы исследования данных, позволяющих формализовать и абстрагировать описание предметной области. Процесс транспортирования информации рассматривается в рамках эталонной семиуровневой модели, известной как модель OSI. Большое внимание уделено протоколам различных уровней, обеспечивающих необходимый уровень стандартизации. Процессы обработки информации излагаются в аспекте поддержки принятия решений с выделением типовых компонентов. Хранение информации представляется с одной стороны, как совокупность моделей концептуального, логического и физического уровней, с другой - как набор методов и способов практической реализации. Большое внимание уделено эргономическим и психологическим факторам при распределении функции между человеком и техническими устройствами в процессе представления и использования информации.

4.1. ИЗВЛЕЧЕНИЕ ИНФОРМАЦИИ

Источниками данных в любой предметной области являются объекты и их свойства, процессы и функции, выполняемые этими объектами или для них. Любая предметная область рассматривается в виде трех представлений (рис. 4.1).

По аналогии с добычей полезных ископаемых процесс извлечения информации направлен на получение ее наибольшей концентрации. В связи с этим процесс извлечения можно представить,

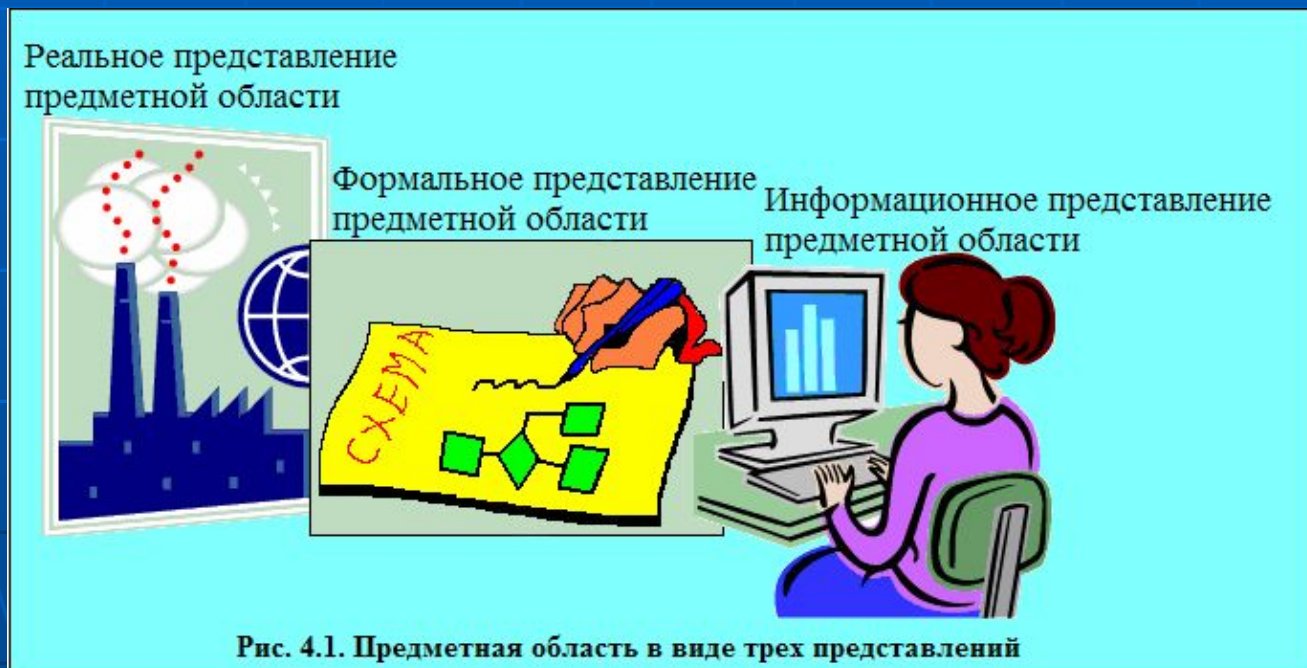


Рис. 4.1. Предметная область в виде трех представлений

как прохождение информации через трехслойный фильтр, в котором осуществляется оценка синтаксической ценности (правильность представления), семантической (смысловой) ценности, прагматической (потребительской) ценности.

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.*

Следует упомянуть неоднородность (разнородность) информационных ресурсов, характерную для многих предметных областей. Одним из путей решения данной проблемы является объектно-ориентированный подход, наиболее распространенный в настоящее время.

Кратко рассмотрим его основные положения.

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.

Экземпляр объекта — это конкретный элемент множества. Например, объектом может являться государственный номер автомобиля, а экземпляром этого объекта — конкретный номер К 173 ПА.

Класс — это множество предметов реального мира, связанных общностью структуры и поведением. Элемент класса — это конкретный элемент данного множества. Например, класс регистрационных номеров автомобиля.

Обобщая эти определения, можно сказать, что объект — это типичный представитель класса, а термины «экземпляр объекта» и «элемент класса» равнозначны. На рис. 4.2 показаны отношения между классами, объектами и предметами реального мира.

Важная особенность объектно-ориентированного подхода связана с понятием инкапсуляции, обозначающим сокрытие данных и методов (действий с объектом) в качестве собственных ресурсов объекта.

Понятия полиморфизма и наследования определяют эволюцию объектно-ориентированной системы, что подразумевает определение новых классов объектов на основе базовых.



Рис. 4.2. Отношения между классами, объектами и предметами реального мира

Полиморфизм интерпретируется как способность объекта принадлежать более чем одному типу.

Наследование выражает возможность определения новых классов на основе существующих с возможностью добавления или переопределения данных и методов.

Для уменьшения избыточности используется процесс обогащения информации, например, при хранении в компьютере списка сотрудников организации иногда достаточно использовать первые 3-4 буквы их фамилий.

Среди методов обогащения информации различают структурное, статистическое, семантическое и прагматическое обогащения.

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.

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 удается 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.