

Лекция 3

Типология БД. Модели представления данных.

Модель данных

Моделью данных называется формализованное описание структуры единиц информации и операций над ними в информационной системе.

Проектирование БД

Концептуальное

Логическое

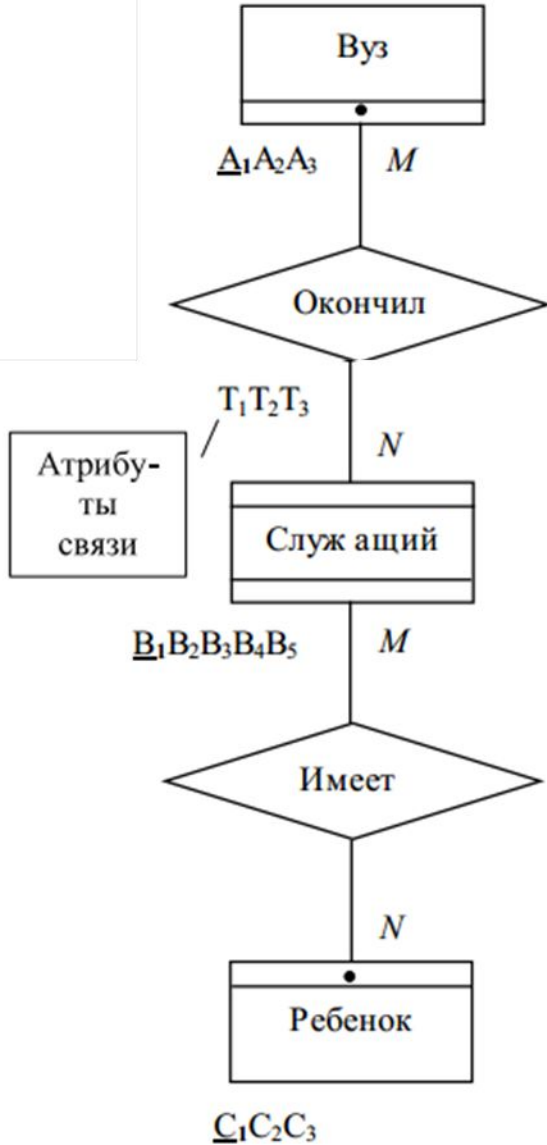
Физическое

Формулирование сущностей, атрибутов и связей

Выбор модели данных и организация данных



Пример схемы составленной на этапе концептуального проектирования



Идентификатор	Семантика	Домен
A ₁	Шифр вуза	C(15)
A ₂	Название	C(30)
A ₃	Город	C(15)

Идентификатор	Семантика	Домен
T ₁	Дата окончания	D
T ₂	Серия и номер диплома	C(30)
T ₃	Специальность	C(15)

Идентификатор	Семантика	Домен
V ₁	Табельный номер	N(6)
V ₂	Ф.И.О.	C(30)
V ₃	Дата рождения	D
V ₄	Дата поступления на предприятие	D
V ₅	Должность	C(5)

Идентификатор	Семантика	Домен
C ₁	Серия и номер свидетельства о рождении	C(30)
C ₂	Ф.И.О.	C(30)
C ₃	Дата рождения	D

Модели данных

Для моделей данных объектом моделирования выступают сами данные, их организация. Это уже не инфологический (семантический), а датологический (синтаксический) аспект рассмотрения будущей базы данных. В этой связи модели данных часто называют синтаксическими.

Синтаксические модели данных.

В модели данных отражаются самые важные аспекты функционирования выделенной предметной области, а второстепенные — игнорируются

Модель данных включает в себя набор понятий для описания:

- ① правил структурирования данных,
- ② процессов динамического изменения данных,
- ③ допустимых состояний взаимосвязанных данных



Компоненты модели данных

Структурная
компонента

Указывает основные структуры данных и правила построения одних структур на основе других

Манипуляционная
компонента

Определяет основные операции над допустимыми структурами данных

Целостная
компонента

Описывает ограничения целостности – специальные требования, которым должны подчиняться структуры данных (например, «вместимость» структуры)



Пример: модель данных алгоритмического языка

Структурная
компонента

Допустимые типы данных. Задаются
декларативными операторами языка

Манипуляционная
компонента

Наборы операций над допустимыми типами
данных. Задаются императивными опера-
торами языка

Целостная
компонента

Диапазоны изменения целых и
вещественных чисел, максимальная длина
строки символов и др.

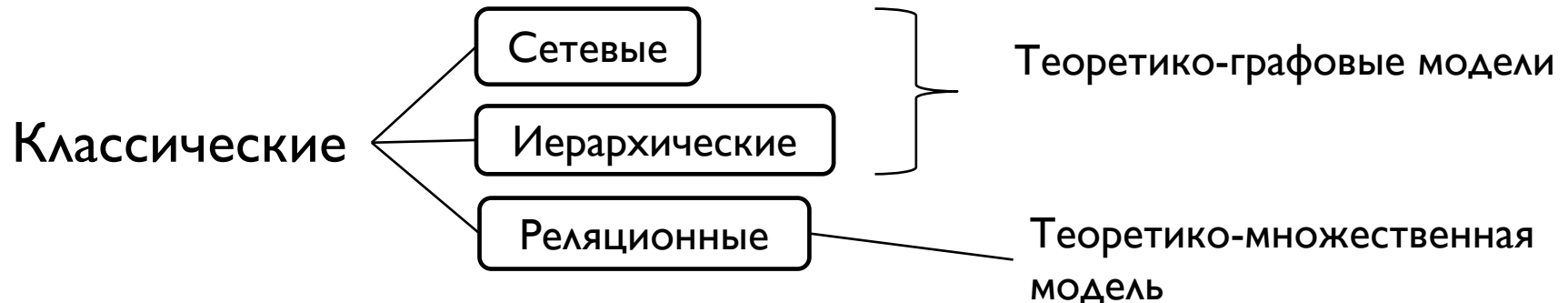


Модели данных

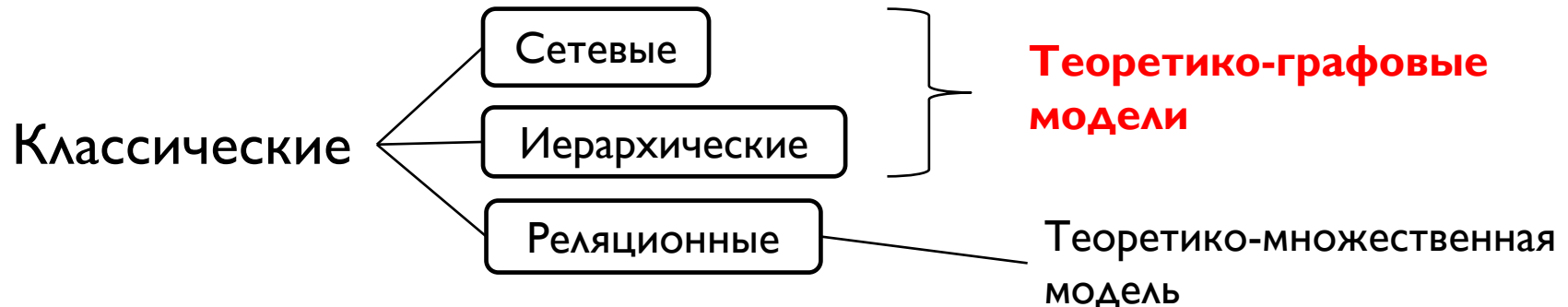
- Каждая СУБД поддерживает ту или иную модель данных.
- По существу модель данных полностью определяет множество конкретных баз данных, которые могут быть созданы средствами этой системы, а также способы модификации состояния БД
- В первую очередь она задает виды «тары», в которую размещаются данные для хранения.
- Не следует отождествлять понятия «модель данных» и «структура данных». Структура данных – лишь одна из трех компонент модели данных



Типы моделей данных



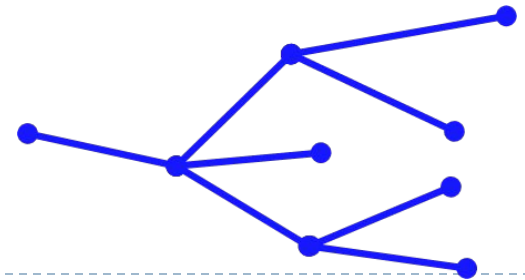
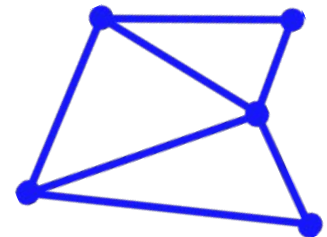
Теоретико-графовые модели данных



В основе **Теоретико-графовых** моделей лежит теория графов. Граф - это способ описания какой-то области реального мира с помощью 2-ух элементов:

- Вершина.
- Дуга.

Исторически эти модели появились раньше, и в настоящий момент они используются реже, чем более современная реляционная модель данных.

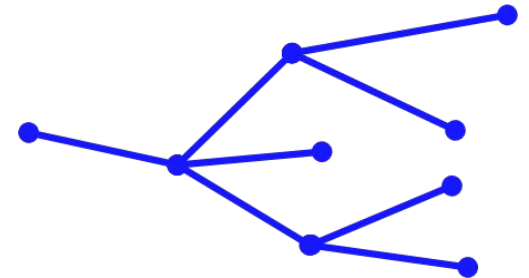


Иерархическая модель

Первые СУБД использовали иерархическую модель данных.

Именно иерархическую модель поддерживала первая из зарегистрированных промышленных СУБД *IMS (Information Management System)* фирмы *IBM*. Эта СУБД была создана для поддержки лунного проекта «Аполлон».

Появление иерархической модели связано с тем, что в реальном мире очень многие связи соответствуют иерархии, когда один объект выступает как основной, а с ним связано множество подчиненных объектов



Иерархическая модель

- В иерархических моделях основная структура представления данных имеет форму дерева.
- На самом высшем уровне иерархии находится только одна вершина, которая называется корнем дерева.³
- Эта вершина имеет связи с вершинами второго уровня, вершины второго уровня имеют связи с вершинами третьего уровня и т.д.
- Связи между вершинами одного уровня отсутствуют. Следовательно, данные в иерархической структуре не равноправны – одни жестко подчинены другим.
- Доступ к информации возможен только по вертикальной схеме, начиная с корня, так как каждый элемент связан только с одним элементом на верхнем уровне и с одним или несколькими на низком.



Иерархическая модель

- **Атрибут** (элемент данных, **поле**) - определяется как наименьшая неделимая единица данных, доступная пользователю.. Обычно каждому элементу при описании базы данных присваивается уникальное имя. По этому имени к нему обращаются при обработке.
- **Запись (сегмент)** - именованная совокупность атрибутов. Использование записей позволяет за одно обращение к базе получить некоторую логически связанную совокупность данных. Именно сегменты изменяются, добавляются и удаляются.
- **Тип записи** —поименованная совокупность входящих в него атрибутов.
- **Экземпляр записи** - конкретная запись с конкретным значением элементов
- **Групповое отношение (связь)** - иерархическое отношение между сегментами двух типов. Родительская запись (владелец группового отношения) называется исходной записью, а дочерние записи (члены группового отношения) - подчиненными. Иерархическая база данных может хранить только такие древовидные структуры

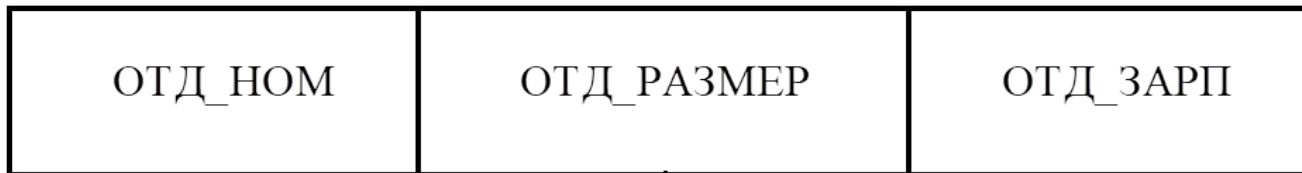
Иерархическая модель

- Записи (сегменты) связываются с помощью отношений в наборы – **корневые деревья**.
- **Корень дерева** – это выделенная запись, выступающая в роли владельца набора.
- Типы записей составляют **тип дерева**
- Экземпляры записей составляют **экземпляр дерева**
- Иерархическая база данных представляет собой упорядоченную совокупность экземпляров деревьев

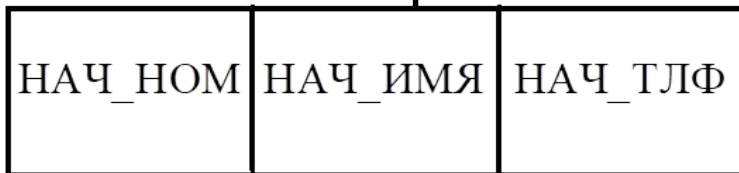


Пример описания типа дерева

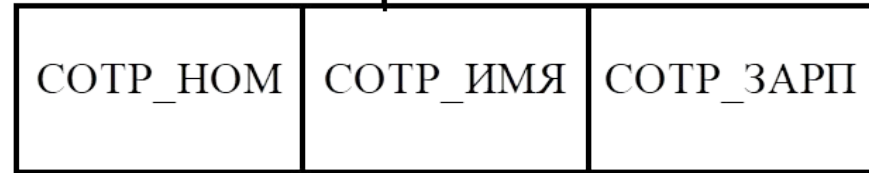
ОТДЕЛ



НАЧАЛЬНИК



СОТРУДНИКИ



Экземпляр дерева

ОТДЕЛ

310	25	1000000
-----	----	---------

НАЧАЛЬНИК

1846	Иванов	3-66
------	--------	------

СОТРУДНИКИ

843	Петров	150000
-----	--------	--------

844	Сидоров	155000
-----	---------	--------



Компоненты модели данных

Структурная
компонента

Указывает основные структуры данных и правила построения одних структур на основе других

Манипуляционная
компонента

Определяет основные операции над допустимыми структурами данных

Целостная
компонента

Описывает ограничения целостности – специальные требования, которым должны подчиняться структуры данных (например, «вместимость» структуры)




Компоненты иерархической модели данных

Структурная
компонента

Основная структура данных – сегмент (запись), построенный из скалярных и составных типов данных. Сегменты связываются в наборы – корневые деревья. Иерархическая база данных представляет собой упорядоченную совокупность экземпляров деревьев.

Манипуляционная
компонента

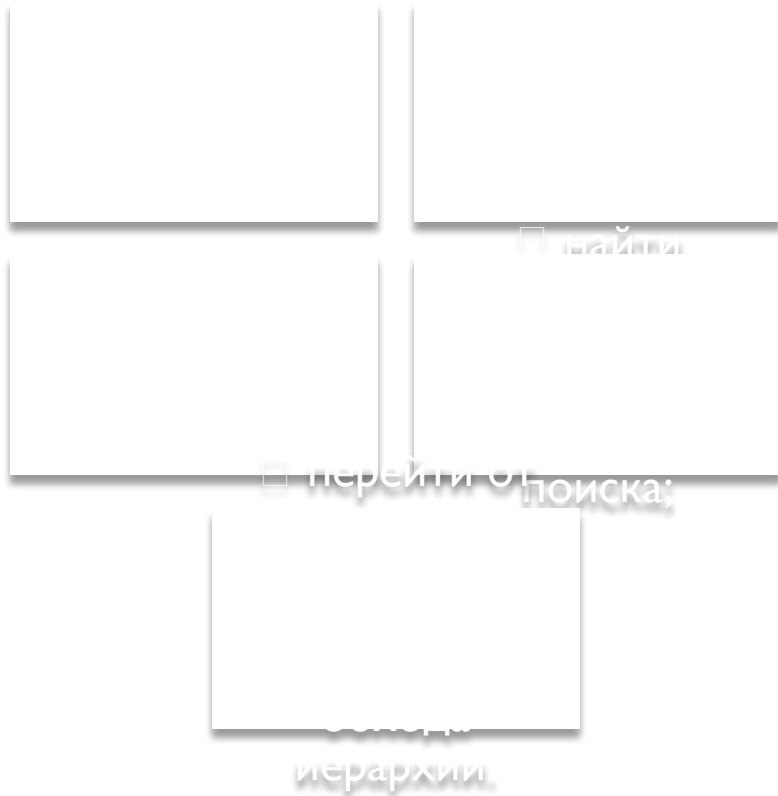
Целостная
компонента



Манипуляционная компонента иерархической модели. Операций манипулирования данными

операторы навигации

операторы
модификации



Компоненты иерархической модели данных

Структурная
компонента

Основная структура данных – сегмент (запись), построенный из скалярных и составных типов данных. Записи связываются в наборы – корневые деревья. Иерархическая база данных представляет собой упорядоченную совокупность экземпляров деревьев.

Манипуляционная
компонента

Операторы навигации и модификации данных

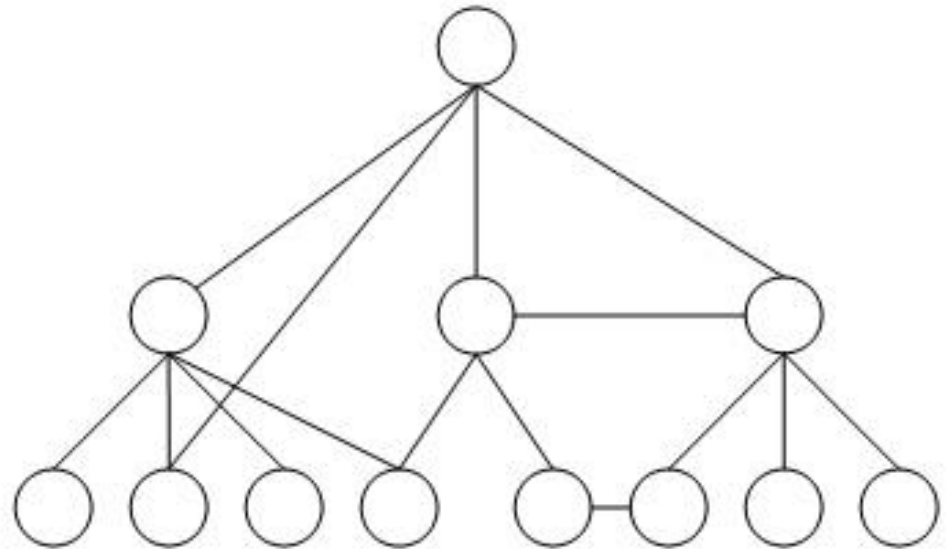
Целостная
компонента

Основное правило:
никакой потомок не может существовать без
своего родителя



Сетевая модель данных

- **Сетевой подход** к организации данных является расширением иерархического.
- Если в иерархической модели запись-потомок должна иметь в точности одного предка, то в сетевой модели она может иметь любое число предков.



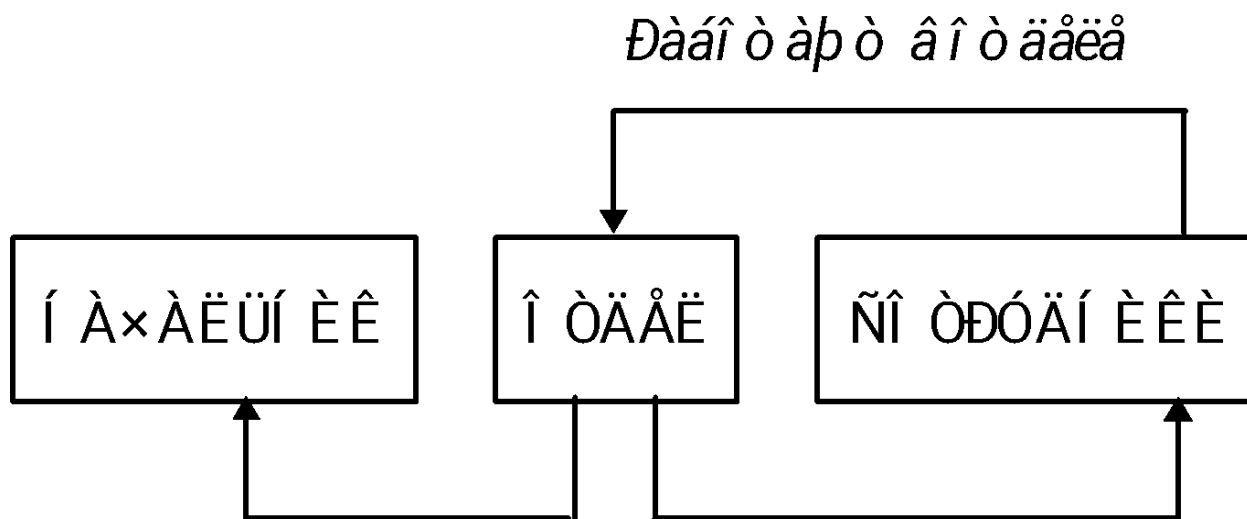
Сетевая модель данных

- У связи теперь появляется имя – в иерархической модели в нем не было необходимости
- **Сетевая база данных** состоит из
 - набора записей (сегментов)
 - и набора связей между записями,

точнее, из набора экземпляров каждого типа записи и набора экземпляров каждого типа связи.



Схема сетевой базы данных



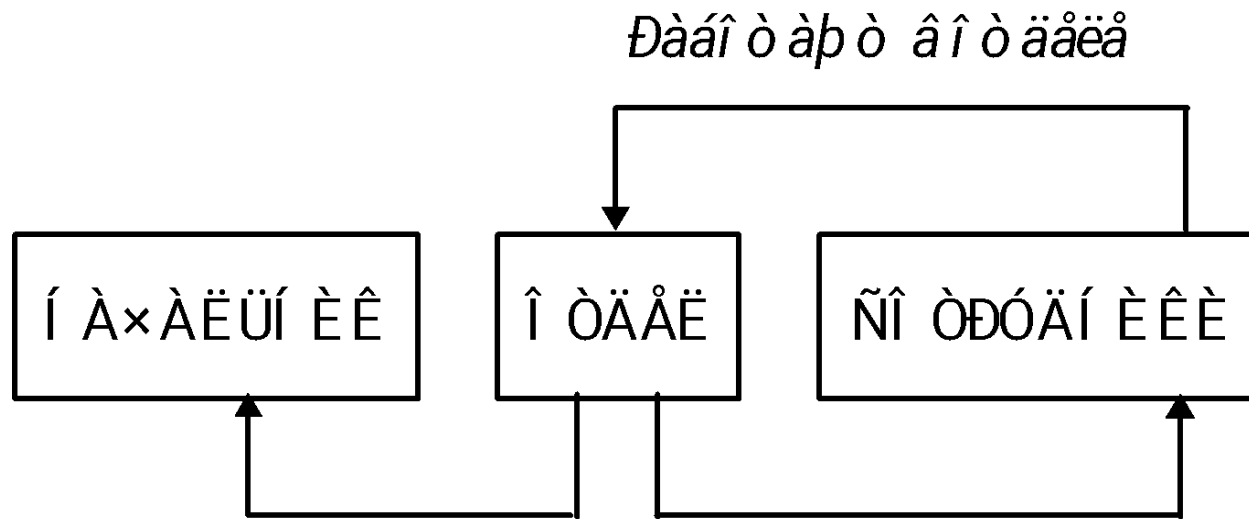
Èì ååò í à÷àëüí èêà Ñî ñò î èò èç ñî ò ðóáí èêî â

Тип связи определяется для двух типов записи – *предка* и *потомка*.

Экземпляр типа связи состоит из одного экземпляра типа записи *предка* и упорядоченного набора экземпляров типа записи *потомка*.



Схема сетевой базы данных



Èì äåò ï à÷àëüí èêà Ñîòðóäíèê òò ðóáîðòàþùèé èêà

Тип связи «Работают в отделе» : предок – сотрудник, потомок – отдел

Экземпляры связи «Работают в отделе»:

Иванов работает в отделе 1

Петров работает в отделе 3 и в отделе 4

Сетевая модель данных

- На формирование типов связи не накладывается особых ограничений. Возможны, например, следующие ситуации:
 - тип записи P может быть в одном типе связи $L1$ потомком, а в другом типе связи $L2$ – предком;
 - тип записи P может быть предком в любом числе типов связи;
 - тип записи P может быть потомком в любом числе типов связи;
 - возможно любое количество типов связи, где P – предок, а C – потомок;
 - типы записи X и Y могут быть предками и потомками в одной связи, и потомками и предками – в другой;
- ▶ предок P и потомок C могут быть одного типа записи.

Сетевая модель данных. Некоторые особенности

- Сетевая модель данных предполагает наличие в ней произвольного количества записей и связей в том числе их различных типов.
- Связь между двумя записями может выражаться произвольным количеством связей.
- В любом наборе может быть только один предок.
- Тип записи может быть предком в одних типах связей и потомком в других типах связей, а также не входить ни в какой тип связей
- При удалении записи-предка удаляются соответствующие связи на экземпляры-потомки, но сами записи-потомки не уничтожаются

Манипуляционная компонента сетевой модели. Операций манипулирования данными

навигационные операции



операторы
модификации



Компоненты сетевой модели данных

Структурная
компонента

Основная структура данных – запись. Записи связываются с помощью связей в наборы – ориентированные графы. Каждый набор имеет выделенную запись – владельца набора. Сетевая база данных – упорядоченная совокупность экземпляров наборов

Манипуляционная
компонента

Операторы навигации и модификации данных

Целостная
компонента

Требуется поддержка целостности по ссылкам,
Например, если удаляется какая-то запись, то должен быть удален и соответствующий экземпляр типа связи.



иерархических и сетевых моделей данных

Достоинства

- 1) простота (для иерархических БД) и высокая гибкость (для сетевых БД) средств представления информации;
- 2) развитые средства управления данными во внешней памяти на низком уровне;
- 3) возможность построения эффективных программ;
- 4) возможность экономии памяти благодаря высокой гибкости структур.

Недостатки

- 1) слишком сложный язык манипулирования данными, которым довольно трудно пользоваться;
- 2) логика перегружена деталями организации доступа к БД;
- 3) фактически требуется знание о физической организации данных;
- 4) прикладные программы зависят от организации данных

Указанных недостатков нет в реляционных базах данных