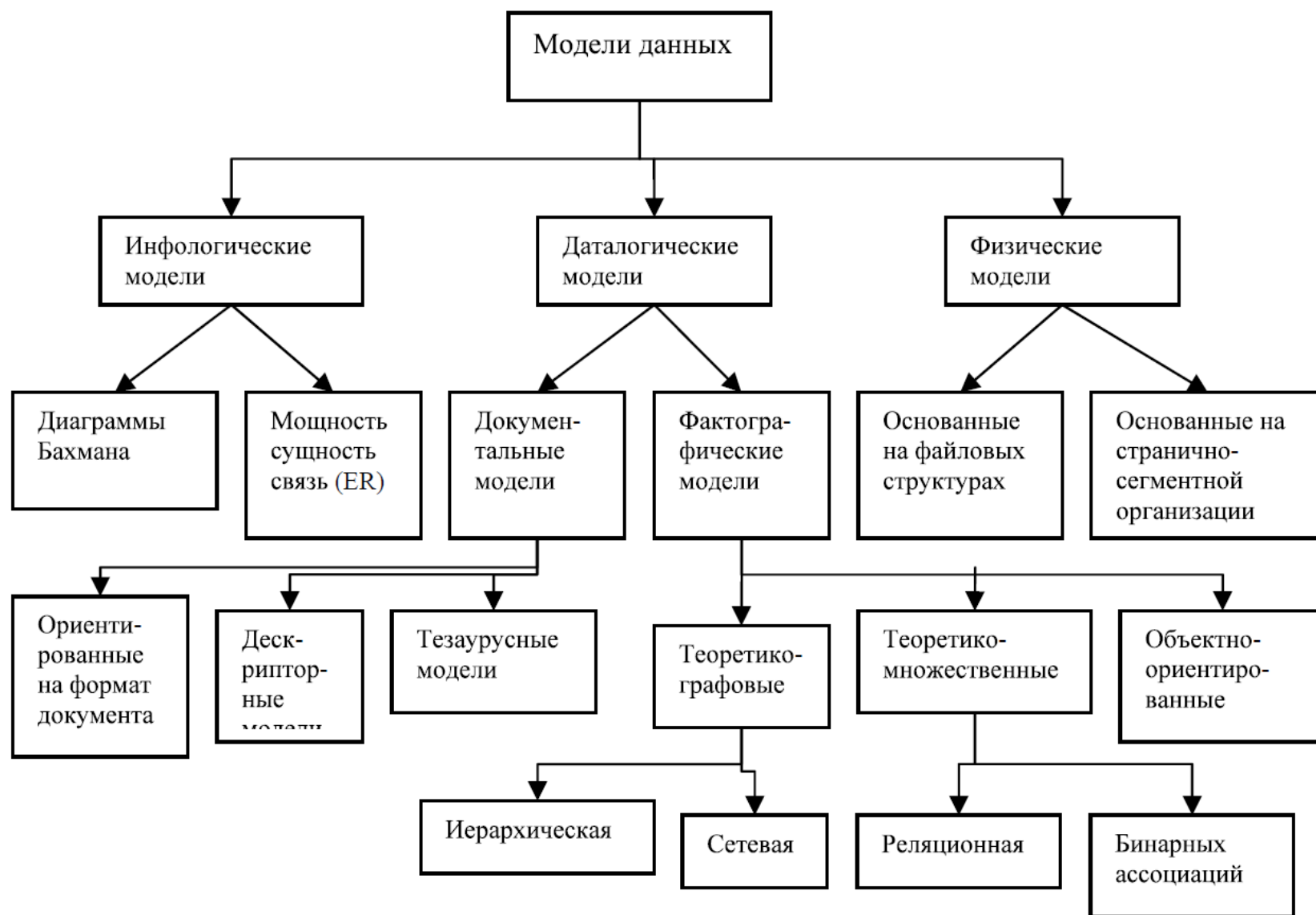


Классификация моделей данных



Инфологические модели используются на ранних стадиях проектирования.

Даталогические модели поддерживаются конкретной СУБД.

Физические модели определяют способы размещения данных в среде хранения и способы доступа к этим данным, которые поддерживаются на физическом уровне.

Классификация файлов, используемых в системах баз данных



Файлы с постоянной длиной

Для файлов с постоянной длиной записи адрес размещения записи с номером K может быть вычислен по формуле

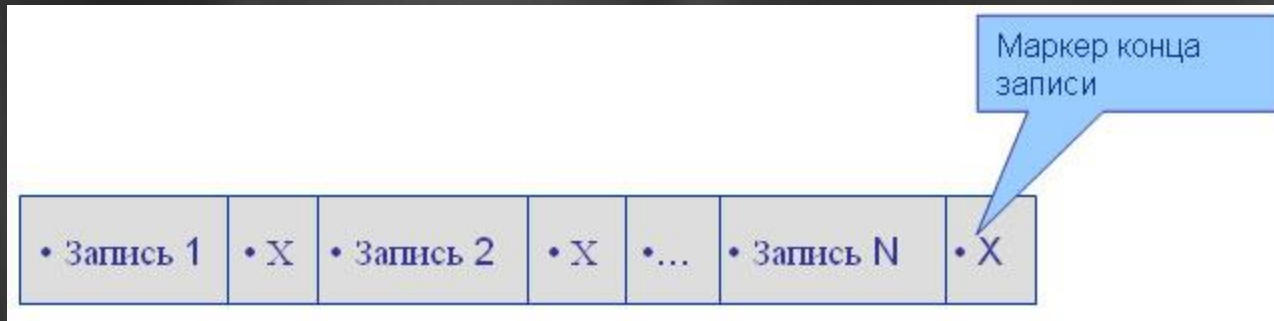
$$BA + (K - 1) * LZ + 1,$$

где BA — базовый адрес, LZ — длина записи.

Файлы с переменной длиной

Файлы с переменной длиной записи всегда являются файлами последовательного доступа. Они могут быть организованы двумя способами:

1. Конец записи отличается специальным маркером.



2. В начале каждой записи записывается ее длина.

LZ1	Запись1	LZ2	Запись2	LZ3	Запись3
-----	---------	-----	---------	-----	---------

Функция ключа

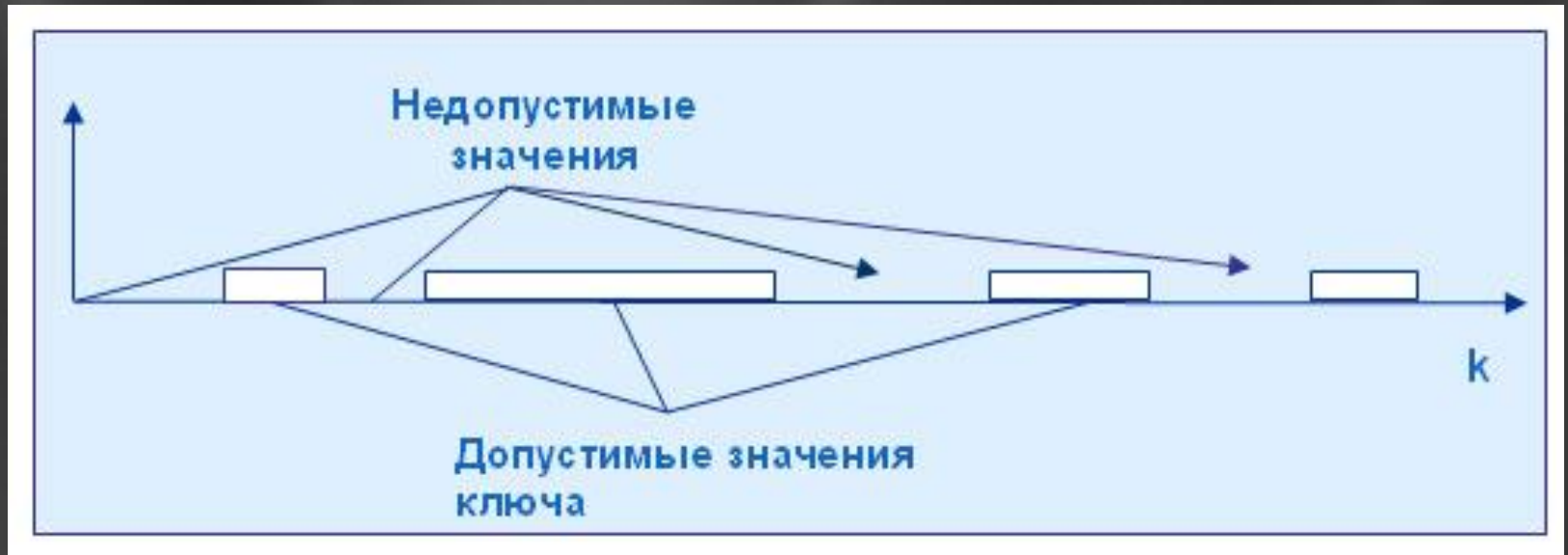
При организации файлов прямого доступа в некоторых очень редких случаях возможно построение функции, которая по значению ключа однозначно вычисляет адрес (номер записи файла):

$$NZ = F(K),$$

где NZ — номер записи, K — значение ключа, $F()$ — функция.

Функция $F()$ при этом должна быть монотонной, чтобы обеспечивать однозначное соответствие.

Неравномерное распределение ключей



В этом случае не удастся построить взаимнооднозначную функцию, либо эта функция будет иметь множество незадействованных значений, которые соответствуют недопустимым значениям ключа. В подобных случаях применяют различные методы хэширования (рандомизации) и создают специальные хэш-функции

Иерархическая модель

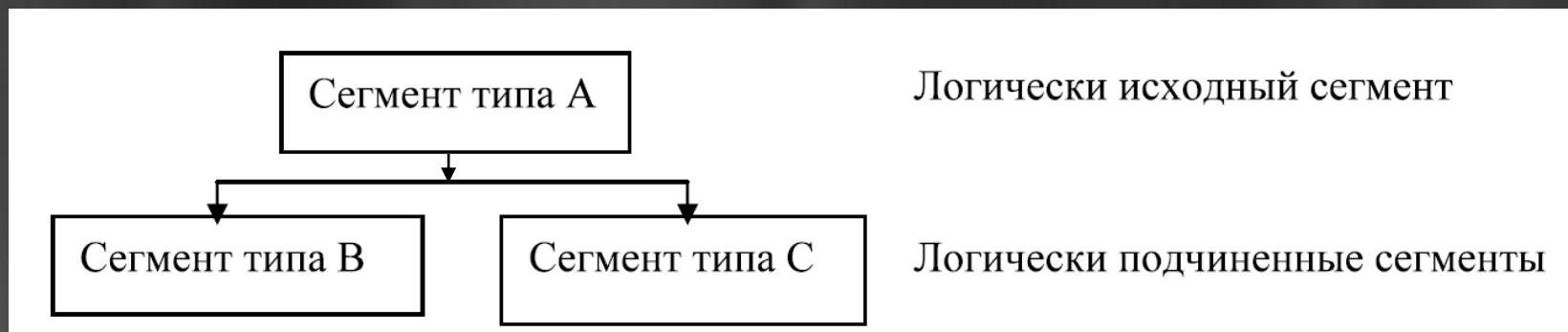
Основные информационные единицы:

- база данных
- сегмент
- поле

Иерархическая модель

Сегмент

- тип сегмента
- экземпляр сегмента



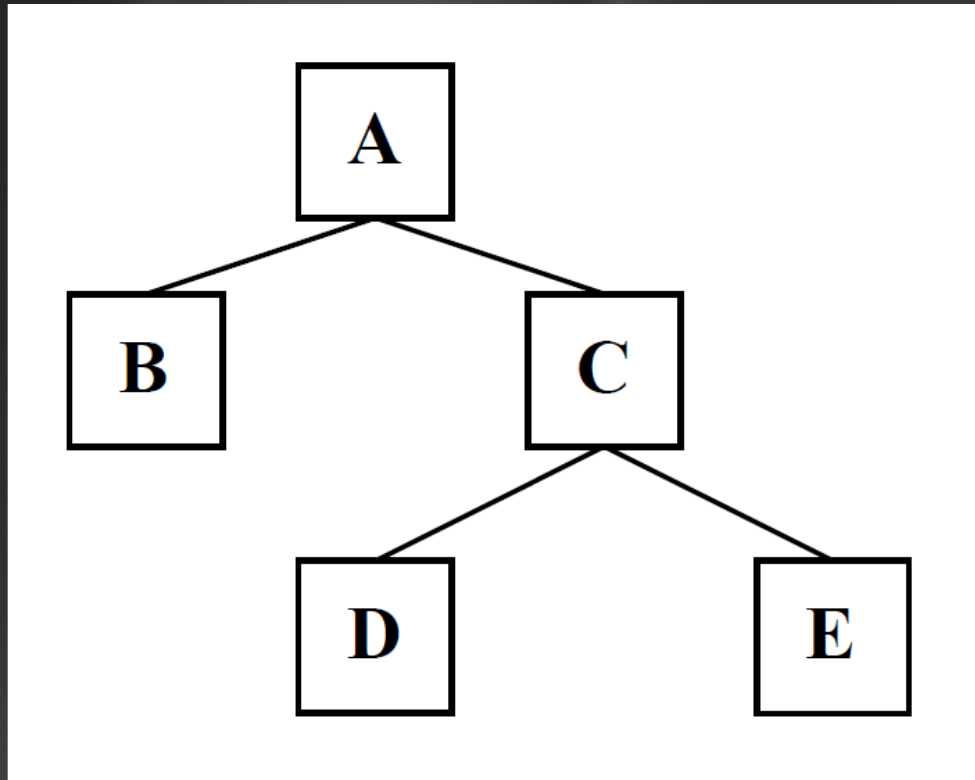
Иерархическая модель

Схема – совокупность отдельных деревьев.
Каждое дерево называется физической базой данных

- только один корневой сегмент
- каждый исходный сегмент может быть связан с произвольным числом подчиненных сегментов
- каждый подчиненный сегмент может быть связан только с одним исходным сегментом

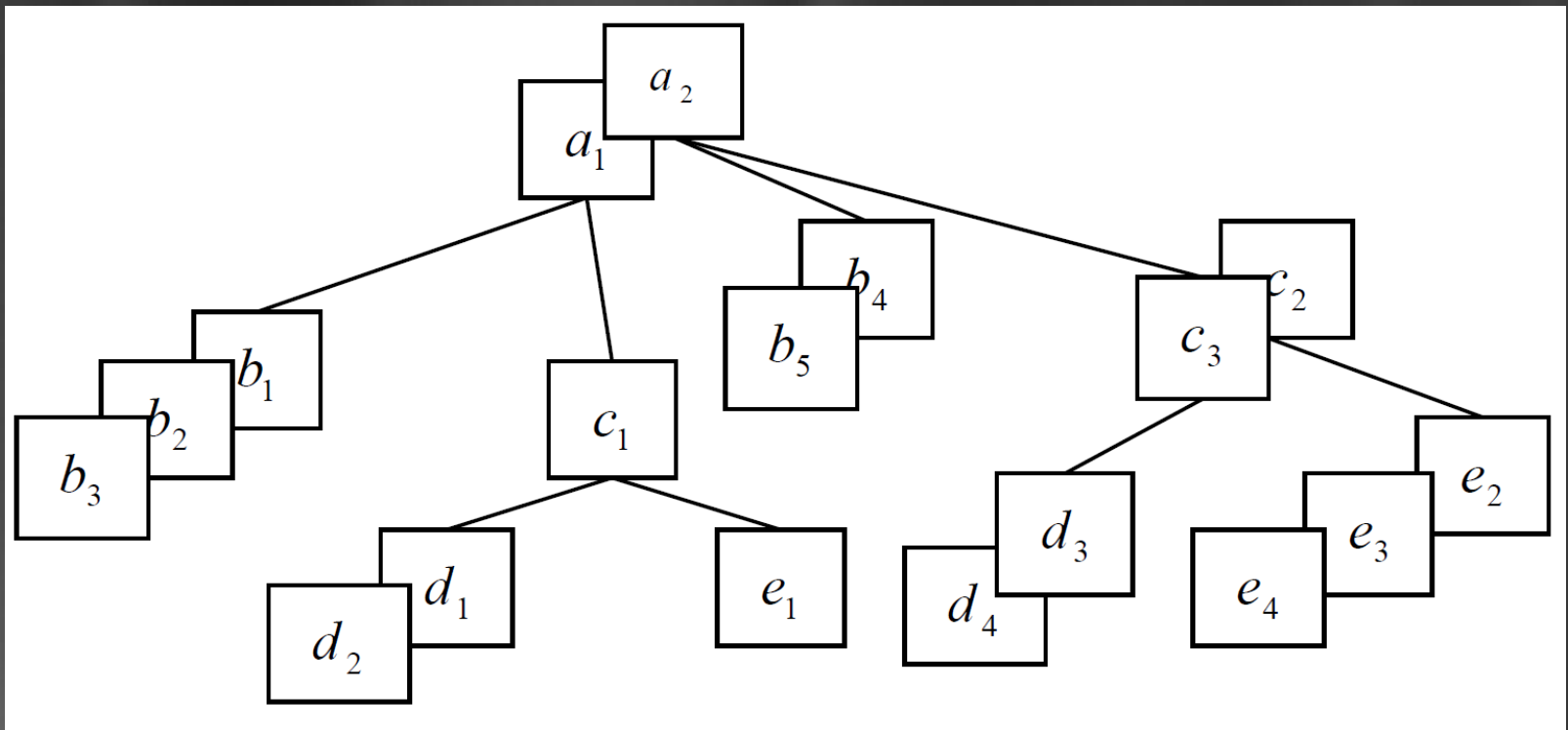
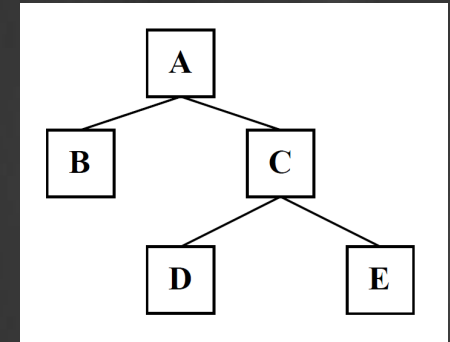
Иерархическая модель

- Иерархический граф



Иерархическая модель

- 2 экземпляра дерева



Иерархическая модель

- Набор всех экземпляров сегментов, подчиненных одному экземпляру корневого сегмента, называется физической записью.

$a_1 b_1 b_2 b_3 c_1 d_1 d_2 e_1$	$a_2 b_4 b_5 c_2 d_3 d_4 e_2 e_3 e_4$
Запись 1	Запись 2

Иерархическая модель

Операции над данными

- **ДОБАВИТЬ** в базу данных новую запись. Для корневой записи обязательно формирование значения ключа.
- **ИЗМЕНИТЬ** значение данных предварительно извлеченной записи. Ключевые данные не должны подвергаться изменениям.
- **УДАЛИТЬ** некоторую запись и все подчиненные ей записи.
- **ИЗВЛЕЧЬ:**
 - извлечь корневую запись по ключевому значению, допускается также последовательный просмотр корневых записей
 - извлечь следующую запись

Иерархическая модель

Достоинства модели

- эффективное использование памяти ЭВМ
- высокая скорость выполнения основных операций над данными
- удобство работы с иерархически упорядоченной информацией
- простота при работе с небольшим объемом данных, так как иерархический принцип соподчиненности понятий является естественным для многих задач

Иерархическая модель

Недостатки модели

- Неуниверсальность. Многие важные варианты взаимосвязи данных невозможно реализовать средствами иерархической модели, или реализация связана с повышением избыточности в базе данных
- Трудность в применении к данным со сложной внутренней взаимосвязью
- Громоздкость такой модели для обработки информации с достаточно сложными логическими связями
- Трудность в понимании ее функционирования обычным пользователем

Иерархическая модель

Ограничения целостности.

Поддерживается только целостность связей между владельцами и членами группового отношения (никакой потомок не может существовать без предка).

Сетевая модель

Базовые объекты:

- элемент данных – минимальная информационная единица
- агрегат данных (вектор и повторяющаяся группа)
- запись – совокупность агрегатов или элементов данных, моделирующая класс объектов реального мира (аналог сегмента)
- набор данных – двухуровневый граф для двух типов записи

Сетевая модель

Агрегат-вектор

Адрес			
Город	Улица	дом	квартира

Агрегат-повторяющаяся группа

Зарплата	
Месяц	Сумма

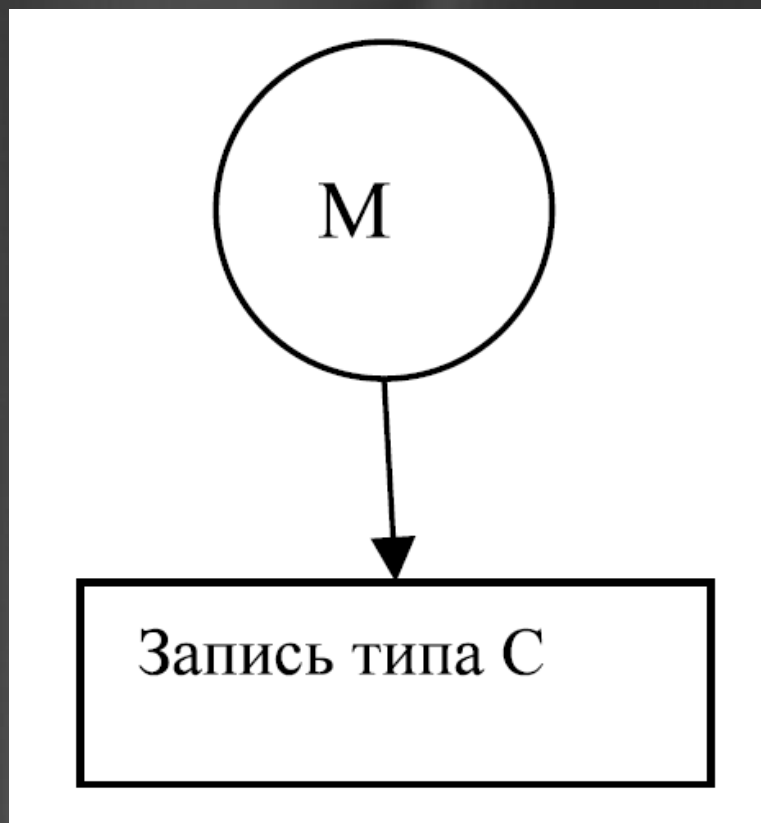
Сетевая модель

Набор



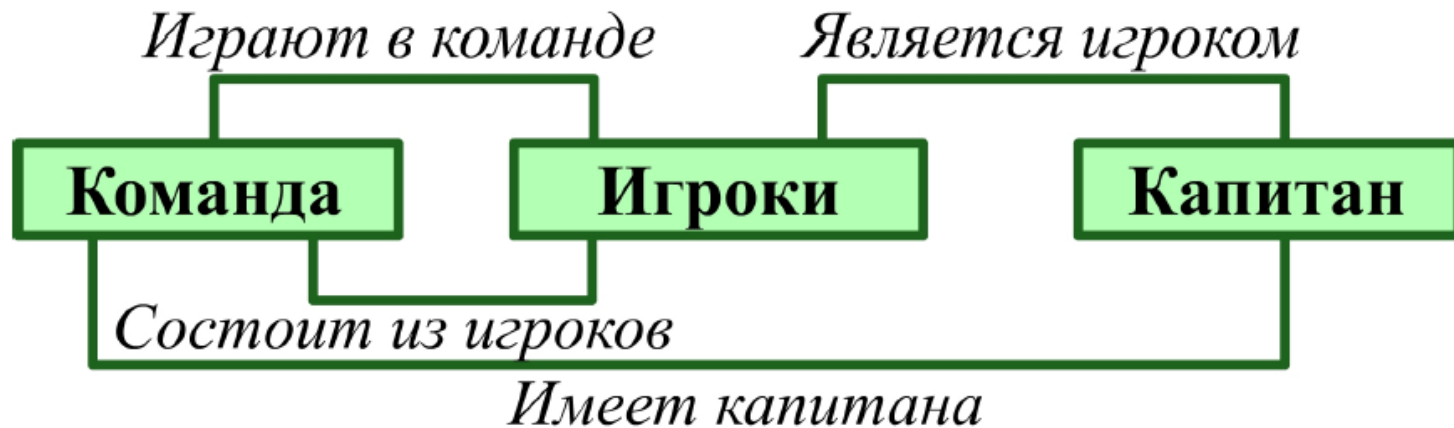
Сетевая модель

Сингулярный набор



Сетевая модель

Пример



Сетевая модель

Преимущества сетевой модели данных:

- Универсальность. Выразительные возможности сетевой модели данных являются наиболее обширными в сравнении с остальными моделями.
- Возможность доступа к данным через значения нескольких отношений (например, через любые основные отношения).

Сетевая модель

Недостатки сетевой модели данных:

- Сложность, т.е. обилие понятий, вариантов их взаимосвязей и особенностей реализации.
- Допустимость только навигационного принципа доступа к данным.