

Основные понятия и определения баз данных

Назначение базы данных – организация информации так, чтобы можно было легко и быстро находить её, чтобы одну и ту же совокупность данных можно было использовать для максимального числа приложений.

Структурирование — это введение соглашений о способах представления данных.

Личное дело № 16493, Сергеев Петр Михайлович, дата рождения 1 января 1976; личное дело № 16593, Петрова Анна Владимировна, дата рождения 15 марта 1975; личное дело № 16693, Анохин Андрей Борисович, дата рождения 14 апреля 1976.

Пример неструктурированных данных

№ личного дела	Фамилия	Имя	Отчество	Дата рождения
16493	Сергеев	Петр	Михайлович	01.01.76
16593	Петрова	Анна	Владимировна	15.03.75
16693	Анохин	Андрей	Борисович	14.04.76

Пример структурированных данных

База данных (БД) — это именованная совокупность структурированных и взаимосвязанных данных, относящихся к одной предметной области.

База данных — именованная совокупность данных, отражающая состояние объектов и их отношений в рассматриваемой предметной области.

Предметная область – это часть
реального мира, определяющая
информационные потребности системы.

Система управления базами данных (СУБД) – это комплекс программных и языковых средств, необходимых для создания баз данных, поддержания их в актуальном состоянии и организации поиска в них необходимой информации.

СУБД – это универсальное прикладное программное средство, предназначенное для создания и ведения (обслуживания) в различных предметных областях баз данных на внешних запоминающих устройствах компьютера, а также доступа к данным и их обработки.

Типовые функции обработки данных в базах данных

1. Добавление и удаление данных.
2. Изменение данных – модификация значений данных в полях записей.
3. Выборка данных – выборка записей из массивов в соответствии с заданными условиями. Выборка осуществляется средствами запросов.
4. Обработка данных из одного массива или взаимосвязанных данных из нескольких массивов.

Основные требования к базам данных

1. Отсутствие дублирования (избыточности) данных, обеспечивающее однократный ввод данных и соответственно простоту корректировки.
2. Целостность и непротиворечивость данных — это такое наполнение базы данными, при котором все записи из разных массивов имеют корректные логические связи с записями других массивов, в случае если такие связи определены в логической структуре БД.

Основные требования к базам данных

3. Возможность многоаспектного доступа — обеспечиваются всевозможные выборки из массивов не дублированной информации и многоцелевое использование одних и тех же данных различными задачами и приложениями пользователя.
4. Защита и восстановление данных при аварийных ситуациях, аппаратных и программных сбоях, ошибках пользователя.
5. Возможность модификации структуры базы данных без повторной загрузки данных.

Модели данных

Модель данных — это совокупность структур данных и операций их обработки.

Модель данных — это метод (принцип) логической организации данных, реализуемый в СУБД.

Каждая конкретная СУБД работает с определенной моделью данных.

**Модели данных различаются по
способу установления связей между
данными:**

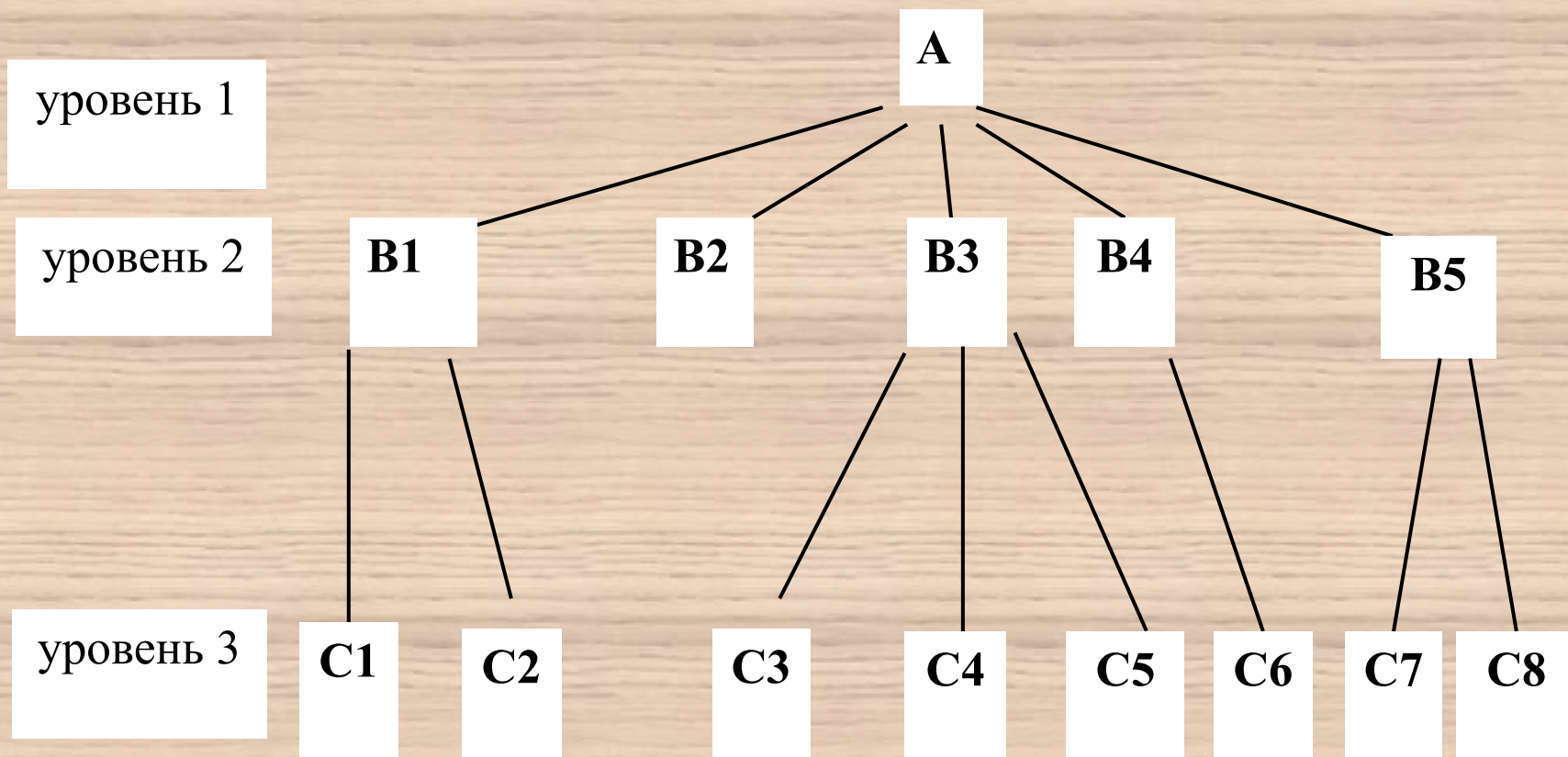
- иерархическая,
- сетевая,
- реляционная.

Иерархическая модель

Иерархическая модель позволяет строить базы данных с древовидной структурой, где каждый узел содержит свой тип данных

Дерево представляет собой иерархию элементов, называемых узлами.

Узел – совокупность атрибутов данных, описывающий некоторый объект.



Требования к ИМД:

- ✓ Каждый узел на более низком уровне связан только с одним узлом, находящимся на более высоком уровне.
- ✓ Иерархическое дерево имеет только одну вершину.
- ✓ К каждому объекту базы данных существует только один путь от вершины.

Достоинства:

- ✓ наличие промышленных СУБД, поддерживающих данную модель;
- ✓ простота описания иерархических структур реального мира,
- ✓ быстрое выполнение запросов.

Недостатки:

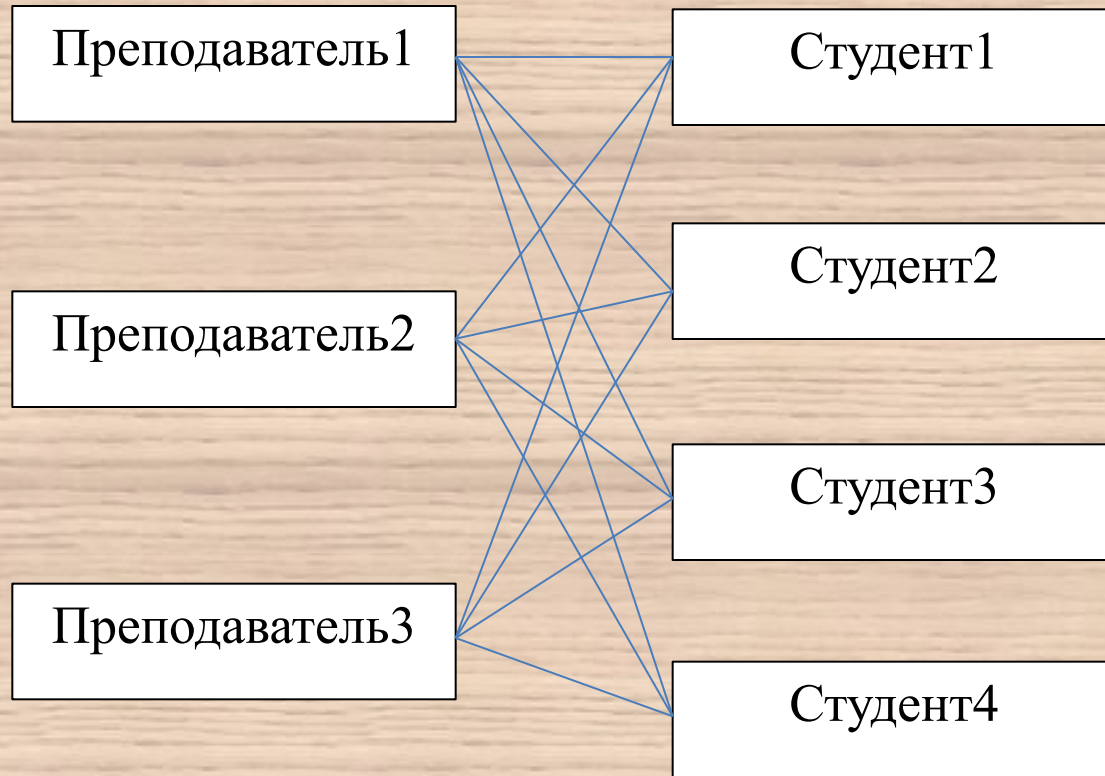
- ✓ иерархия в значительной степени усложняет операции включения информации о новых объектах в БД и удаления устаревшей;
- ✓ доступ к любому узлу возможен только через корневой.

Институт (специальность, название, директор)



Сетевая модель

Каждый объект в сетевой модели может быть связан с любым другим объектом.



Сетевая структура модели БД

Недостатки сетевой модели: сложность и
возможная потеря независимости данных при
реорганизации БД.

Реляционная модель

Реляционная модель данных ориентирована на организацию данных в виде двумерных таблиц.

Реляционная база данных – это совокупность отношений, содержащих всю информацию, которая должна храниться в БД.

Реляционная база данных – это множество взаимосвязанных двумерных реляционных таблиц.

Основные элементы реляционных баз данных

Реляционная таблица – это двумерная таблица, в которой содержатся сведения об одной сущности.

Структура реляционной таблицы определяется составом и последовательностью полей, соответствующих ее столбцам.

Содержание реляционной таблицы заключено в строках таблицы.

Структурные компоненты таблицы БД



№ личного дела	Фамилия	Имя	Отчество	Дата рождения	Группа
16493	Сергеев	Петр	Михайлович	01.01.76	111
16593	Петрова	Анна	Владимировна	15.03.75	112
16693	Анохин	Андрей	Борисович	14.04.76	113

Свойства реляционных таблиц:

1. каждая таблица состоит из однотипных строк и имеет уникальное имя;
2. один элемент таблицы — один элемент данных;
3. все поля таблицы содержат однородные по типу данные (целочисленный, числовой, текстовый, и т.д.);
4. каждое поле имеет уникальное имя;

Свойства реляционных таблиц:

5. число полей задается при создании таблицы;
6. записи имеют фиксированное число полей и значений;
7. порядок записей в отношении может быть произвольным;

Свойства реляционных таблиц:

8. строки таблицы обязательно отличаются друг от друга хотя бы единственным значением;
9. количество записей в отношении не ограничено.

В реляционной таблице каждый столбец есть **домен** (его альтернативное название **поле**), а совокупность элементов каждой строки – **кортеж** (или **запись**).

Сущность – это объект любой природы, данные о котором хранятся в БД.

Пример. ВУЗ, группа в ВУЗе, каждый студент в группе – это сущности.

Сущностями могут быть не только материальные предметы.

Тип сущности – набор однородных личностей, предметов, событий или идей, выступающих как единое целое.

Экземпляр сущности – конкретная вещь в наборе.

Атрибут – это поименованная характеристика сущности, т.е. некоторый показатель, который характеризует некий объект и принимает для конкретного экземпляра объекта некоторое числовое, текстовое или иное значение.

Связь – это ассоциирование двух и более сущностей.

Типы связей

1. один к одному (1:1),
2. один ко многим (1:M),
3. многие ко многим (M:M).

1. Связь один к одному означает, что каждому экземпляру первого объекта (А) может соответствовать только один экземпляр второго объекта (В) и, на оборот, каждому экземпляру второго объекта (В) может соответствовать только один экземпляр первого объекта (А).

2. Связь один ко многим означает, что каждому экземпляру одного объекта (А) может соответствовать несколько экземпляров другого объекта (В), а каждому экземпляру второго объекта (В) может соответствовать только один экземпляр первого объекта (А).

3. Связь многие ко многим означает, что каждому экземпляру одного объекта (А) могут соответствовать несколько экземпляров второго объекта (В) и, наоборот, каждому экземпляру второго объекта (В) могут соответствовать тоже несколько экземпляров первого объекта (А).

Основные типы связей в реляционных таблицах

1. 1:1 - отношение 1:1 предполагает, что каждой записи одной таблицы соответствует одна запись в другой;
2. 1:M - отношение 1:M предполагает, что каждой записи первой таблицы соответствует много записей во второй, но каждой записи второй таблицы соответствует только одна запись в первой.

Первичный ключ – это поле или совокупность полей (атрибутов), значение которых однозначно идентифицирует каждую запись.

Если таблица удовлетворяет требованию уникальности первичного ключа, она называется **отношением**.

Строка заголовков называется **схемой отношения**.

Пример:

СТУДЕНТ (ФАМИЛИЯ, ИМЯ, ОТЧЕСТВО,
ФАКУЛЬТЕТ, КУРС, ГРУППА)

Если первичный ключ состоит из одного атрибута, он называется **простым**, если из нескольких - **составным** первичным ключом.

По значению первичного ключа может быть найден единственный экземпляр строки.

Вторичный ключ (ВК) — это такой атрибут, значение которого может повторяться в нескольких записях таблицы, т.е. он не является уникальным.

В реляционной модели все таблицы должны быть преобразованы в отношения.

Отношения реляционной модели связаны между собой.

Связи поддерживаются внешними ключами.

Внешний ключ – это поле (совокупность полей), значение которых однозначно характеризует значения первичного ключа другого отношения (таблицы).

Свойства первичного ключа

- ✓ уникальность – в таблице может быть назначен только один первичный ключ, у составного ключа поля могут повторяться, но не все;
- ✓ избыточность – не должно быть полей, которые, будучи удаленными из первичного ключа, не нарушат его уникальность;
- ✓ в состав первичного ключа не должны входить поля типа, комментарий и графическое.