

ГАПОУ "Камский строительный колледж
имени Е.Н.Батенчука"

Учебная дисциплина
СИСТЕМЫ УПРАВЛЕНИЯ БАЗЫ ДАННЫХ

Лекция 1. Определение и классификация ИС

Руководитель занятия:
преподаватель спецдисциплин

Еремина И.И.

концепция баз данных

Стержневые идеи современных информационных технологий базируются на концепции *баз данных*.

Согласно этой концепции, основой информационных технологий являются *данные*, которые должны быть организованы в базы данных в целях адекватного отображения изменяющегося реального мира и удовлетворения информационных потребностей пользователей.

Основные понятия теории БД

Одним из важнейших понятий в теории баз данных является понятие *информации*. **Под информацией понимаются любые сведения о каком-либо событии, процессе, объекте.**

Данные — это информация, представленная в определенном виде, позволяющем автоматизировать ее сбор, хранение и дальнейшую обработку человеком или информационным средством.

Для компьютерных технологий данные — это информация в дискретном, фиксированном виде, удобная для хранения, обработки на ЭВМ, а также для передачи по каналам связи.

База данных (БД) — именованная совокупность данных, отражающая состояние объектов и их отношений в рассматриваемой предметной области,
или иначе БД — это совокупность взаимосвязанных данных при такой минимальной избыточности, которая допускает их использование оптимальным образом для одного или нескольких приложений в определенной предметной области. БД состоит из множества связанных файлов.

Система управления базами данных (СУБД)

— совокупность языковых и программных средств, предназначенных для создания, ведения и совместного использования БД многими пользователями.

Автоматизированная информационная система (АИС) — это система, реализующая автоматизированный сбор, обработку, манипулирование данными, функционирующая на основе ЭВМ и других технических средств и включающая соответствующее программное обеспечение (ПО) и персонал.

В дальнейшем в этом качестве будет использоваться термин ***информационная система (ИС)***, который подразумевает понятие автоматизированная.

Каждая ИС в зависимости от ее назначения имеет дело с той или иной частью реального мира, которую принято называть ***предметной областью (ПрО) системы.***

Выявление ПрО — это необходимый начальный этап разработки любой ИС. Именно на этом этапе определяются информационные потребности всей совокупности пользователей будущей системы, которые, в свою очередь, предопределяют содержание ее базы данных.

Банк данных (Бнд) является разновидностью ИС.

Бнд — это система специальным образом организованных данных: баз данных, программных, технических, языковых, организационно-методических средств, предназначенных для обеспечения централизованного накопления и коллективного многоцелевого использования данных.

Под **задачами обработки данных** обычно понимается специальный класс решаемых на ЭВМ задач, связанных с видом, хранением, сортировкой, отбором по заданному условию и группировкой записей однородной структуры.

Отдельные программы или комплекс программ, реализующие автоматизацию решения прикладных задач обработки данных, называются приложениями.

Приложения, созданные средствами СУБД, относят к *приложениям СУБД*. Приложения, созданные вне среды СУБД с помощью систем программирования, использующих средства доступа к БД, к примеру, Delphi или Visual Studio, называют *внешними приложениями*.

Современное состояние технологий баз данных

Кратко сформулируем основные современные принципы организации баз данных.

1. Значительная часть современных СУБД способна работать на компьютерах различной архитектуры под управлением разных операционных систем.
2. Подавляющее большинство современных СУБД обеспечивают поддержку полной реляционной модели данных, обеспечивая целостность категорий и целостность на уровне ссылок.
3. Современные СУБД для определения данных и манипуляции ими опираются на принятые стандарты в области языков, а при обмене данными между различными СУБД базируются на существующих технологиях по обмену информацией.

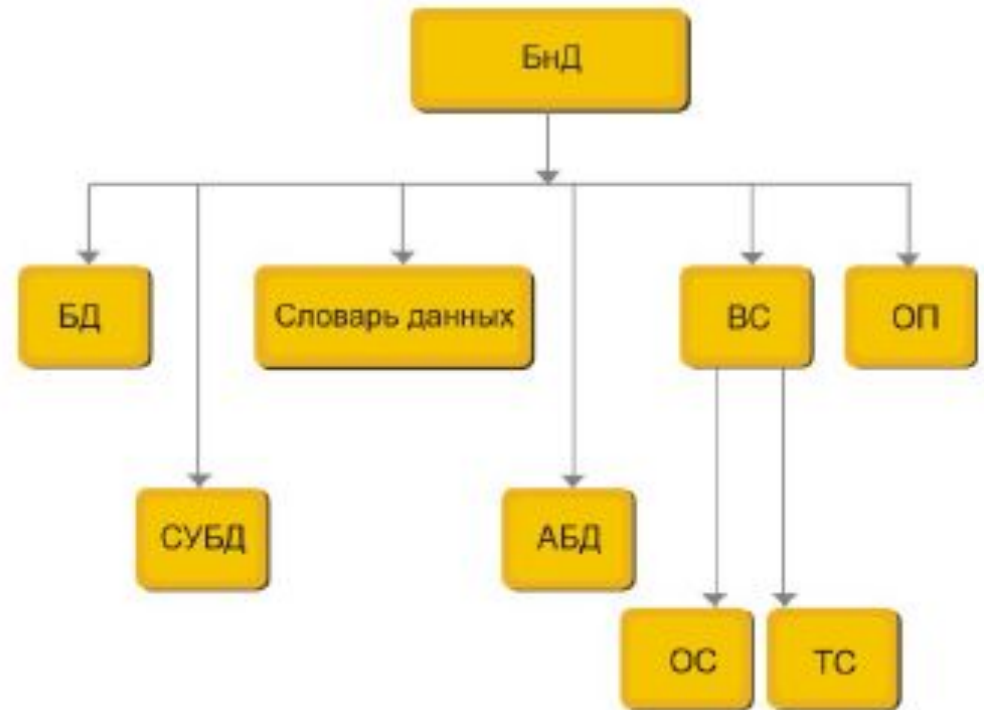
Современное состояние технологий баз данных

4. Многие существующие СУБД относятся к так называемым сетевым СУБД, которые предназначены для поддержки многопользовательского режима работы с базой данных и поддержки возможности децентрализованного хранения данных.
5. Такие СУБД имеют развитые средства администрирования баз данных и средства защиты хранимой в них информации.
6. Подобные СУБД имеют средства подключения клиентских приложений.
7. Современные СУБД характеризуются опытами применения концепции фундаментальной идеи объектно-ориентированного подхода, способствующей повышению уровня абстракции баз данных, являющейся перспективным этапом на пути развития технологий баз

Информационные системы, созданные средствами технологии баз данных, иногда принято называть *банками данных (БнД)*.

БнД включает в себя:

- технические средства;
- одну или несколько БД;
- СУБД;
- словарь или каталог данных;
- администратора;
- вычислительную систему;
- обслуживающий персонал.



Автоматизированные информационные системы (АИС)

В развитии АИС наметились два поколения:

1. **ИС, базирующиеся на автономных файлах.** Это системы с простой архитектурой и ограниченным кругом возможностей. Они состоят из набора автономных файлов и комплекса прикладных программ, предназначенных для обработки этих файлов и выдачи документов. Такие системы имеют ряд серьезных недостатков, ограничивающих их широкое применение: высокую избыточность данных, сложность ведения и совместной обработки файлов, зависимость программ от данных и др.

2. **Банки данных.** Это системы с высокой степенью интеграции данных и автоматизации управления ими. Они ориентированы на коллективное пользование и в основном лишены недостатков, присущих АИС 1-го поколения.

Пользователи ИС

- *Администраторы системы,* отвечающие за ее эксплуатацию;
- *Прикладные программисты,* разрабатывающие прикладные программы для решения различных задач;
- *Конечные пользователи,* составляющие наиболее многочисленную группу потребителей информации.

Классификации ИС по типу хранимых данных

1. **Документальные информационно-поисковые системы (ДИСП)** предназначены для хранения и обработки документальных данных-адресов хранения документов, наименований, описаний и рефератов, а также текстов документов. Такие данные представляются в неструктурированном виде. Примером ДИСП являются библиотечные и библиографические АИС.
2. **Фактографические информационно-поисковые системы (ФИПС)** хранят и обрабатывают структурированные данные в виде чисел и текстов. Над такими данными можно выполнять различные операции.

Классификации ИС по характеру обработки данных

1. **Информационно – справочные системы (ИСС)**, запросно-ответные или просто справочные, которые выполняют поиск и вывод информации без ее обработки.
2. **Автоматизированные информационные системы обработки данных**, сочетающие в себе ИИС с системой обработки данных . Обработка найденных данных выполняется комплексом прикладных

Классификации ИС по степени

интеграции данных и

автоматизации управления ими

- 1. На автономных файлах (первое поколение)** принцип интеграции практически **не использовался**, а уровень автоматизации управления файлами был сравнительно низким. Такие системы применяются и в настоящее время, они эффективны в случае узкого, специализированного использования небольшим кругом лиц.
- 2. Банки данных (БНД)** с **большой** степенью интеграции данных.

Классификации ИС по степени распределенности данных

- 1. Локальные*, размещенные на одной ЭВМ.
- 2. Распределенные*, функционирующие в среде вычислительной сети, распределенные по ее узлам (серверам и рабочим станциям).

Требования к банкам данных (БнД)

- **Адекватность информации состоянию предметной области.** БнД являются информационной моделью предметной области. Хранимая в нем информация должна полно и точно отображать ее объекты, их свойства и отношения между ними. Отступление от принципа адекватности делает систему бесполезной и даже опасной, недопустимой для использования.
- **Надежность функционирования** – одно из важнейших требований, предъявляемых к любой системе.
- **Быстродействие и производительность.** Эти два близких друг другу требования отражают временные потребности пользователей. Первое определяется временем ответа (реакции) системы на запрос, исчисляемым с момента ввода запроса до момента начала выдачи найденных данных. Второе требование определяется количеством запросов, выполняемых в единицу времени.

Требования к банкам данных (БнД)

- **Быстрота и удобство использования.** Это требование предъявляется со стороны всех категорий пользователей, особенно конечных.
- **Массовость использования.** Необходимо обеспечить коллективный доступ пользователей, при котором они могут одновременно и независимо обращаться к базам данных для получения необходимых сведений.
- **Защита информации** от случайных искажений, уничтожения, а также от преднамеренных, несанкционированных действий пользователя.
- **Возможность расширения.** Архитектура системы должна допускать расширение ее возможностей путем модификации или замены существующих программных модулей либо добавления новых компонентов, а также путем реорганизации информационных массивов.

Принципы построения банков данных

- **Принцип интеграции данных**, объединении отдельных, взаимно несвязанных данных в единое целое, в роли которого выступает база данных, результате чего пользователю и его прикладным программам все данные представляются единым информационным массивом..
- **Принцип централизации управления данными**, передача всех функций управления данными единому комплексу управляющих программ – СУБД.

Оба принципа отражают суть банка данных: интеграция является основой организации БД, централизация управления – основой организации и функционирования системы управления базами данных (СУБД).

Принципы построения банков

данных

- **Принцип целостности данных** отражает требование адекватности хранимой в БД информации состоянию предметной области.
- **Под независимостью данных** понимают независимость прикладных программ от хранимых данных, при которой любые изменения в организации данных не требуют коррекции этих программ. Одним из путей достижения независимости является введение дополнительных уровней абстрагирования данных (**принцип многоуровневости**)
- **Неизбыточность** - это состояние данных, когда каждое из них присутствует в базе данных в единственном экземпляре. Избыточность может иметь место как на логическом уровне, когда в структуре данных повторяются одни и те же типы данных, так и на физическом уровне, когда данные хранятся в двух или

Принципы построения банков

данных

- Под **непротиворечивостью** понимается смысловое соответствие между данными; это состояние БД, при котором хранимые в ней данные не противоречат друг другу. Различают два аспекта непротиворечивости: **смысловое соответствие** разнотипных данных и **идентичность** (равенство) дублирующих данных.
- **Принцип связности данных** заключается в том, что данные в БД взаимосвязаны, и связи отражают отношения между объектами предметной области. Множество связей и множество типов данных образуют логическую структуру данных. Наличие связей между записями в БД позволяет уменьшить избыточность, упростить и ускорить поиск
- данных.

Состав банка данных

Состав БНД определяется исходя из возлагаемых на систему функций.

Основными функциями являются:

- ❑ **Хранение информации** и организация ее защиты;
- ❑ **Периодическое изменение хранимых данных** (обновление, удаление, добавление);
- ❑ **Поиск и отбор данных по запросам** пользователей и прикладных программ;
- ❑ **Обработка** найденных записей и вывод результатов в заданной форме.

Необходим комплекс программ, которые бы обеспечивали автоматизацию всех операций, связанных с решением этих задач - **СУБД**.

СУБД

СУБД – это совокупность программ и языковых средств, предназначенных для создания, ведения и использования БД. СУБД представляет собой пакет прикладных программ, расширяющих возможности операционной системы по обработке БД.

- **Ядро** – управляющая программа СУБД
- **Набор обрабатывающих программ**: трансляторов с языков описания данных, языков запросов и языков программирования, редакторов, отладчиков
- **Инструментальные средства и средства организации доступа** к БД. Обработка найденных системой данных, вычисления, формирование выходных документов по заданной форме выполняются с помощью прикладных программ.

Обеспечение банков данных



Классификация банков данных

(по типу используемой модели)

Хранимые в базе данные имеют определенную логическую структуру – иными словами, описываются некоторой моделью представления данных (моделью данных), поддерживаемой СУБД. К числу классических относят следующие модели данных:

- иерархическую;
- сетевую;
- реляционную.

В последнее время на практике активно используют

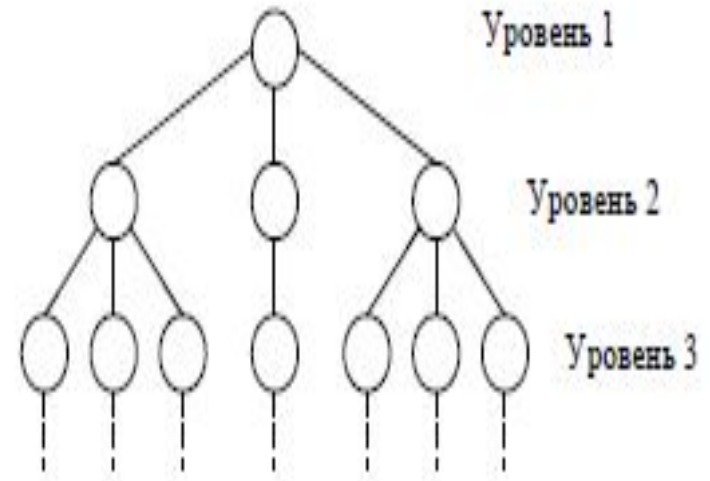
- постреляционная;
- многомерная;
- объектно-ориентированная

МОДЕЛИ ДАННЫХ

Иерархическая модель

Иерархическая структура представляет совокупность элементов, связанных между собой по определенным правилам. Графическим способом представления иерархической структуры является дерево .

Дерево представляет собой иерархию элементов, называемых узлами. Под элементами понимается совокупность атрибутов, описывающих объекты. В модели имеется корневой узел (корень дерева), который находится на самом верхнем уровне и не имеет узлов, стоящих выше него. У одного дерева



МОДЕЛИ ДАННЫХ

Иерархическая модель

К *достоинствам* иерархической модели данных относятся эффективное использование памяти ЭВМ и неплохие показатели времени выполнения операций над данными.

Недостатком иерархической модели является ее громоздкость для обработки информации с достаточно сложными логическими связями.

На иерархической модели данных основано сравнительно ограниченное количество СУБД, в числе которых можно назвать зарубежные системы IMS, PC/Focus, Teradata, Data Base

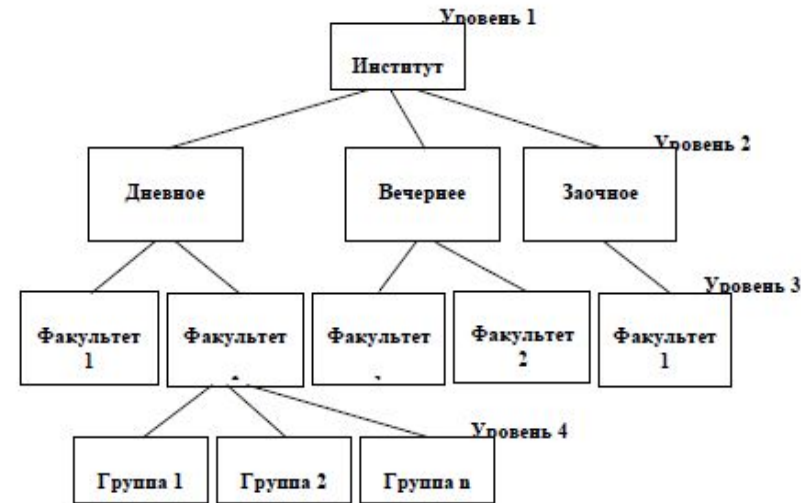


Рис. 2.2-1. Пример иерархической структуры

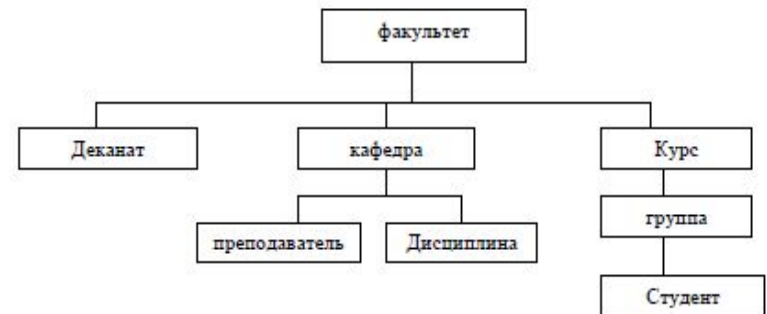


Рис. 2.2-2. Пример иерархической структуры

МОДЕЛИ ДАННЫХ

Сетевая модель данных

Отличие сетевой структуры от иерархической заключается в том, что каждый элемент в сетевой структуре может быть связан с любым другим элементом.

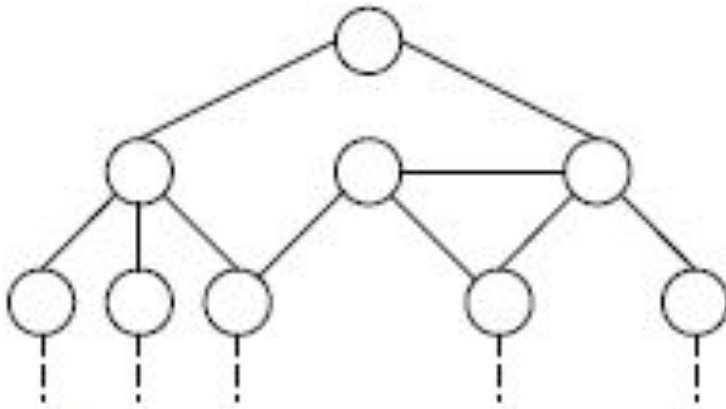


Рис. 2.3. Графическое изображение сетевой структуры

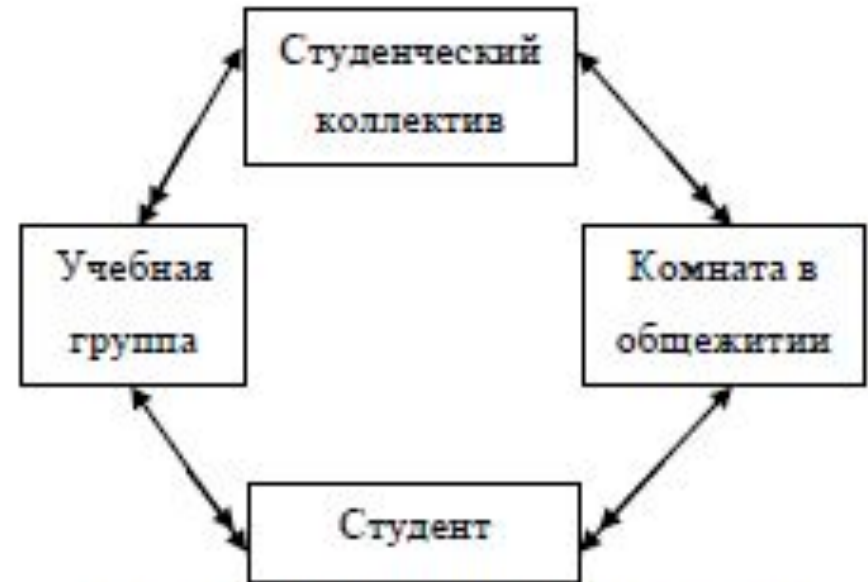


Рис. 2.4. Пример взаимосвязей между элементами сетевой структуры

МОДЕЛИ ДАННЫХ

Сетевая модель данных

Достоинством сетевой модели данных является возможность эффективной реализации по показателям затрат памяти и оперативности.

Недостатком сетевой модели данных являются высокая сложность и жесткость схемы БД, построенной на ее основе.

Наиболее известными сетевыми СУБД являются IDMS, db_VistaIII, СЕТЬ, СЕТОР и КОМПАС.

МОДЕЛИ ДАННЫХ

Реляционная модель

данных

Реляционная модель данных была предложена Е.Ф. Коддом, известным исследователем в области баз данных, в 1969 году, когда он был сотрудником фирмы IBM. Впервые основные концепции этой модели были опубликованы в 1970.

Реляционная база данных представляет собой хранилище данных, организованных в виде двумерных таблиц. Любая таблица реляционной базы данных состоит из строк (называемых также записями) и столбцов (называемых также полями).

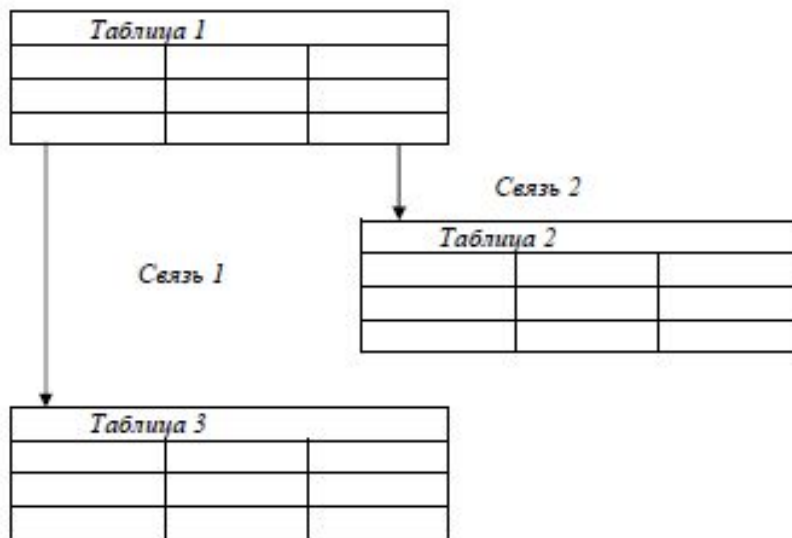
Строки таблицы содержат сведения о представленных в ней фактах (или документах, или людях, одним словом, - об однотипных объектах). На пересечении столбца и строки находятся конкретные значения содержащихся в таблице данных.

МОДЕЛИ ДАННЫХ

Реляционная модель

Данные в таблицах удовлетворяют следующим принципам:

- 1) Каждое значение, содержащееся на пересечении строки и столбца, должно быть атомарным.
- 2) Значения данных в одном и том же столбце должны принадлежать к одному и тому же типу, доступному для использования в данной СУБД.
- 3) Каждая запись в таблице уникальна, то есть в таблице не существует двух записей с полностью совпадающим набором значений ее полей.
- 4) Каждое поле имеет уникальное имя.
- 5) Последовательность полей в



МОДЕЛИ ДАННЫХ

Реляционная модель

данных

Поле или комбинацию полей, значения которых однозначно идентифицируют каждую запись таблицы, называют **ВОЗМОЖНЫМ КЛЮЧОМ** (или просто **КЛЮЧОМ**).

Если таблица имеет более одного возможного ключа, тогда один ключ выделяют в качестве **первичного**. Первичный ключ любой таблицы обязан содержать уникальные непустые значения для каждой строки.

Поле, указывающее на запись в другой таблице, связанную с данной записью, называется **внешним ключом**. Иначе говоря, внешний ключ – это поле или набор полей, чьи значения совпадают с имеющимися значениями первичного ключа другой таблицы.

МОДЕЛИ ДАННЫХ

Реляционная модель

данных

Подобное взаимоотношение между таблицами называется **связью**. Группа связанных таблиц называется **схемой базы данных**. Информация о таблицах, их полях, первичных и внешних ключах, а также иных объектах базы данных, называется **метаданными**.

Достоинство реляционной модели данных заключается в простоте, понятности и удобстве физической реализации на ЭВМ. Именно простота и понятность для пользователя явились основной причиной ее широкого использования.

К основным *недостаткам реляционной модели* относятся отсутствие стандартных средств идентификации отдельных записей и сложность описания иерархических и сетевых связей.

Примерами зарубежных реляционных СУБД для ПЭВМ являются: DB2, Paradox, FoxPro, Access, Clarion, Ingres, Oracle.

К отечественным СУБД реляционного типа относятся

Постреляционная модель данных

Классическая реляционная модель предполагает неделимость данных, хранящихся в полях записей таблиц. **Постреляционная модель представляет собой расширенную реляционную модель,** снимающую ограничение неделимости данных. Модель допускает многозначные поля – поля, значения которых состоят из подзначений. Набор значений многозначных полей считается самостоятельной таблицей, встроенной в основную таблицу.

Накладные		Накладные-товары		
N накладной	Покупатель	N накладной	Товар	Количество
0373	8723	0373	Сыр	3
8374	8232	0373	Рыба	2
7364	8723	8374	Лимонад	1
		8374	Сок	6
		8374	Печенье	2
		7364	Йогурт	1

б)

Накладные			
N накладной	Покупатель	Товар	Количество
0373	8723	Сыр	3
		Рыба	2
8374	8232	Лимонад	1
		Сок	6
		Печенье	2
7364	8723	Йогурт	1

Рис. 2.6. Структуры данных реляционной (а) и постреляционной (б) моделей

МОДЕЛИ ДАННЫХ

Постреляционная модель данных

Достоинством постреляционной модели является возможность представления совокупности связанных реляционных таблиц одной постреляционной таблицей. Это обеспечивает высокую наглядность представления информации и повышение эффективности ее обработки.

Недостатком постреляционной модели является сложность решения проблемы обеспечения целостности и непротиворечивости хранимых данных.

Рассмотренная постреляционная модель данных поддерживается СУБД uniVers. К числу других СУБД, основанных на постреляционной модели данных, относятся также системы Bubba и Dasdb.

МОДЕЛИ ДАННЫХ

Многомерная модель

данных

Многомерный подход к представлению данных появился практически одновременно с реляционным, но интерес к многомерным СУБД появился с середины 90-х годов. Толчком послужила в 1993 году статья Э. Кодда. В ней были сформулированы 12 основных требований к системам класса OLAP (OnLine Analytical Processing –оперативная аналитическая обработка), важнейшие из которых связаны с возможностями концептуального представления и обработки многомерных данных.

Реляционные СУБД предназначались для информационных систем оперативной (транзакционной) обработки информации и в этой области весьма эффективны. В системах аналитической обработки они показали себя несколько неповоротливыми и недостаточно гибкими. Более эффективными здесь оказываются многомерные СУБД.

Многомерные СУБД являются узкоспециализированными СУБД, предназначенными для интерактивной аналитической обработки информации.

Основные понятия, используемые в этих СУБД:

агрегируемость, историчность и прогнозируемость.

МОДЕЛИ ДАННЫХ

Многомерная модель данных

Основные понятия многомерных моделей данных: **измерение и ячейка**.

Измерение – это множество однотипных данных, образующих одну из граней гиперкуба. В многомерной модели измерения играют роль индексов, служащих для идентификации конкретных значений в ячейках гиперкуба.

Ячейка – это поле, значение которого однозначно определяется фиксированным набором измерений. Тип поля чаще всего определен как цифровой. **Значения некоторой ячейки может быть переменной либо формулой.**

Модель	Месяц	Объем
Жигули	июнь	12
Жигули	июль	24
Жигули	август	5
Москвич	июнь	2
Москвич	июль	18
Волга	июль	19

Модель	Июнь	Июль	Август
Жигули	12	24	5
Москвич	2	18	No
Волга	No	19	No

МОДЕЛИ ДАННЫХ

Объектно-ориентированная модель данных (ООМ)

В ООМ при представлении данных имеется возможность идентифицировать отдельные записи базы данных. Между записями и функциями их обработки устанавливаются взаимосвязи с помощью механизмов, подобных средствам в объектно-ориентированных языках программирования.

Стандартизированная ООМ описана в рекомендациях стандарта ODMG-93 (Object Database Management Group – группа управления объектно-ориентированными базами данных).

Рассмотрим упрощенную ООМ БД. Структура графически представима **в виде дерева**, **узлами** которого являются **объекты**. Свойства объектов описываются некоторым стандартным типом или типом, конструируемым пользователем (определяется как class). Значение свойства типа class есть объект, являющийся экземпляром соответствующего класса. Каждый **объект-экземпляр** класса считается потомком объекта, в котором он определен как свойство. Объект-экземпляр класса принадлежит своему классу и имеет **одного родителя**. Родовые отношения в БД образуют связную иерархию объектов.

МОДЕЛИ ДАННЫХ

Объектно-ориентированная модель данных (ООМ)

Логическая структура объектно-ориентированной БД внешне похожа на структуру иерархической БД. Основное различие между ними состоит в методах манипулирования данными (инкапсуляция, наследование и полиморфизм).

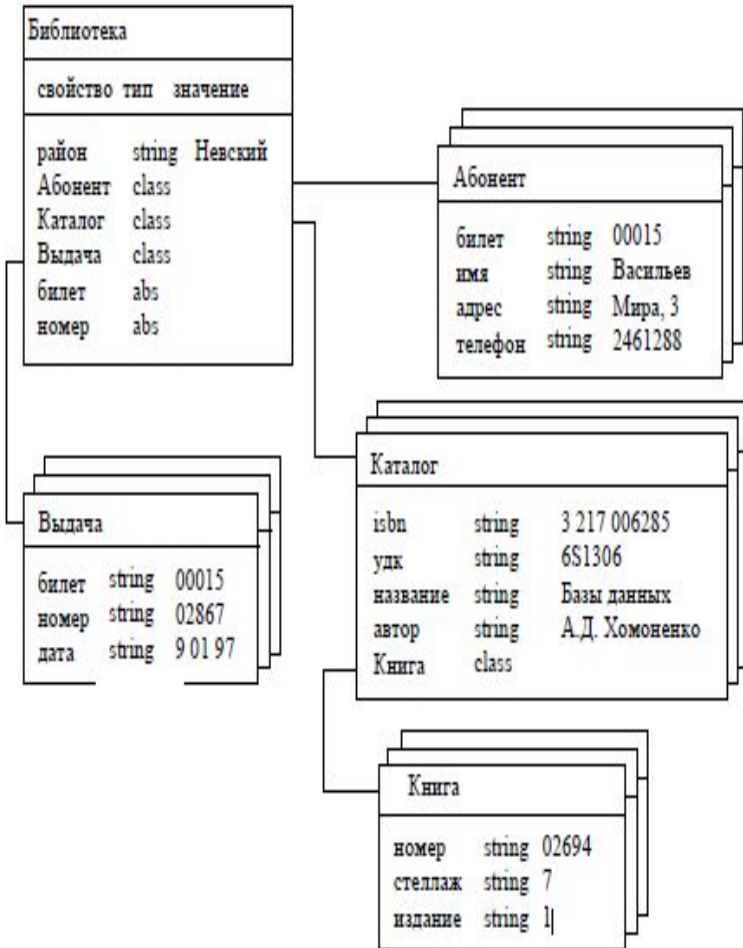
Инкапсуляция ограничивает область видимости имени свойства пределами того объекта, в котором оно определено.

Наследование, наоборот, распространяет область видимости свойства на всех потомков объекта.

Полиморфизм в объектно-ориентированных языках программирования означает способность одного и того же программного кода работать с разнотипными данными. Другими словами, он означает допустимость в объектах разных типов иметь методы (процедуры или функции) с одинаковыми именами.

МОДЕЛИ ДАННЫХ

Объектно-ориентированная модель данных (ООМ)



Основным **достоинством** ООМ данных в сравнении с реляционной является возможность отображения информации о сложных взаимосвязях объектов. Объектно-ориентированная модель данных позволяет идентифицировать отдельную запись базы данных и определять функции их обработки.

Недостатками ООМ являются высокая понятийная сложность, неудобство обработки данных и низкая скорость выполнения запросов.

К объектно-ориентированным СУБД относятся POET, Jasmine, Versant, O2, ODB-Jupiter, Iris, Orion.

Тесты для самоконтроля

1. Под элементами в иерархической структуре

понимают:

- а) совокупность атрибутов, описывающих объекты;
- б) данные, организованные в виде двумерных таблиц;
- в) поля, значения которых однозначно определяются фиксированным набором измерений.

ОТВЕТ:

2. Каждый элемент может быть связан с любым другим элементом:

- а) в иерархической модели;
- б) в сетевой модели;
- в) в реляционной модели.

ОТВЕТ:

3. В таблицах реляционной базы данных:

- а) существенна последовательность полей и несущественна последовательность записей;
- б) несущественна и последовательность полей, и последовательность записей;
- в) несущественна последовательность полей и существенна последовательность записей.

ОТВЕТ:

4. В информационных системах оперативной обработки информации эффективны:

- а) реляционные СУБД;
- б) многомерные СУБД;
- в) иерархические СУБД.

ОТВЕТ:

5. Многомерность многомерной модели

данных означает:

- а) многомерность визуализации данных;
- б) наличие многозначных полей;
- в) многомерное логическое представление структуры информации.

ОТВЕТ:

6. Возможность отображения информации о сложных взаимосвязях объектов имеется:

- а) в реляционной модели;
- б) в постреляционной модели;
- в) в объектно-ориентированной модели.

ОТВЕТ:

**7. Комбинативно-полные,
значения которых
однозначно
идентифицируют каждую
запись таблицы, называют:**

- а) ключом таблицы;
- б) внешним ключом таблицы;
- в) первичным ключом таблицы.

ОТВЕТ:

8. Неделимость данных – обязательный принцип:

- а) постреляционной модели данных;
- б) реляционной модели данных;
- в) объектно-ориентированной модели.

Ответ:

9. В информационных системах аналитической обработки информации эффективны:

- а) реляционные СУБД;
- б) многомерные СУБД.

ОТВЕТ: