1. 数据库系统概述
2. 数据库地位

数据库技术产生于二十世纪六十年代末

数据库技术是信息系统的核心和基础

1. **数据Data**

数据是数据库中存储的基本对象，是描述事物的符号记录

数据可为文字、图形、图像、声音

数据的特点：**数据与其语义是不可分的**

例如：175不是数据，要说明175代表什么，张三175com高、李四175斤重

1. **数据库Database**

数据库是**长期储存在计算机内**、**有组织的、可共享**的**大量数据的集合**（存储数据的仓库）

数据库的基本特征：数据按一定的数据模型组织、描述和储存；可被不同用户共享；冗余度较小；数据独立性较高；易扩展

1. **数据库管理系统Database management system，DBMS**

数据库管理系统是位于用户（DBA操作管理数据库人员）与操作系统之间的一层数据管理**软件**，是基础软件，是一个大型复杂的软件系统（用于对数据库做操作、管理数据库的系统）

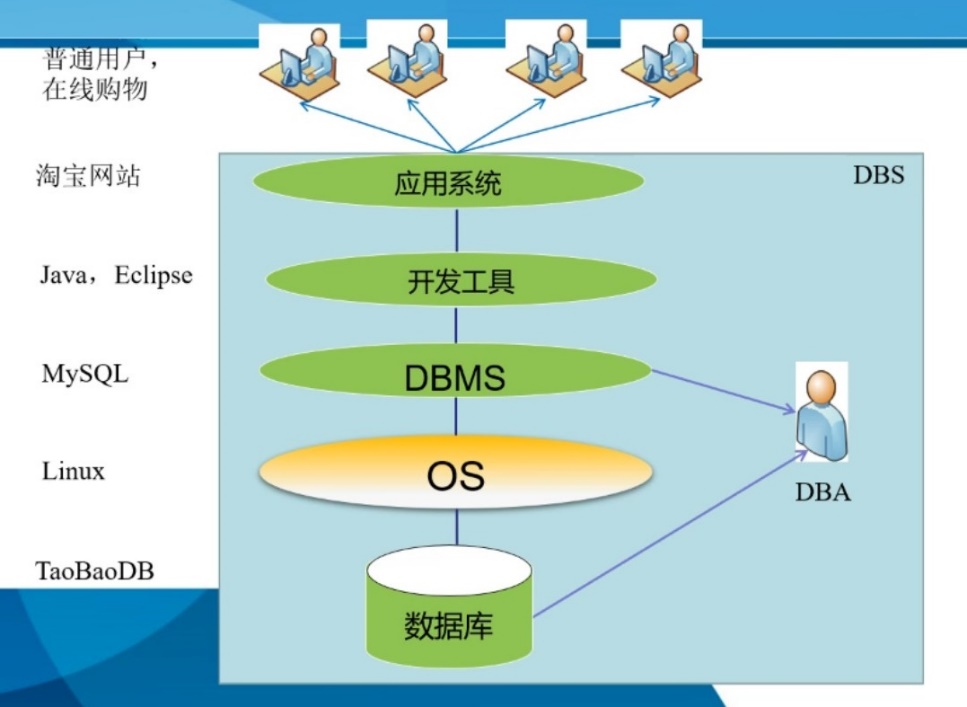
例如：Oracle、SQL server、MySQL、SYBASE等

DBMS主要功能：

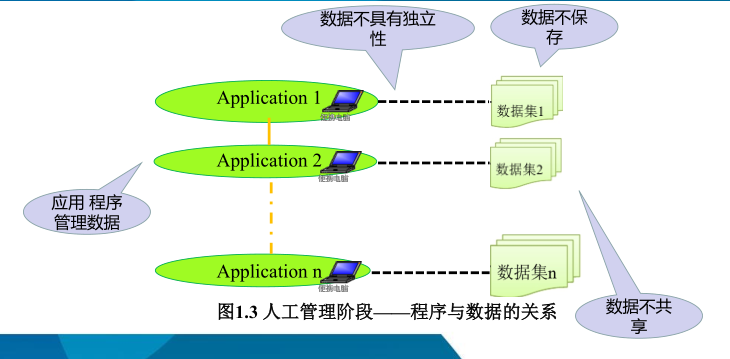
1. 数据定义功能 提供数据定义语言（DDL）、定义数据库中的数据对象
2. 数据组织、存储和管理
3. 数据操纵功能 提供数据操纵语言（DML）、实现对数据库的基本操作（增删改查）
4. 数据库的事物管理和运行管理 发生故障后的恢复
5. 数据库的建立和维护功能
6. **数据库系统Database System，DBS**

数据库系统：在计算机系统中引入数据库后的系统

DBS = DB + DBMS（及其应用开发工具） +应用程序（支付宝、雪梨等） +DBA（数据库管理员）



1. 数据管理技术的产生和发展
2. 数据管理：对数据进行分类、组织、编码、存储、检索和维护，数据处理的中心问题
3. **人工管理阶段**



特点：数据的管理者：用户（程序员），数据不保存

数据面向的对象：某一应用程序

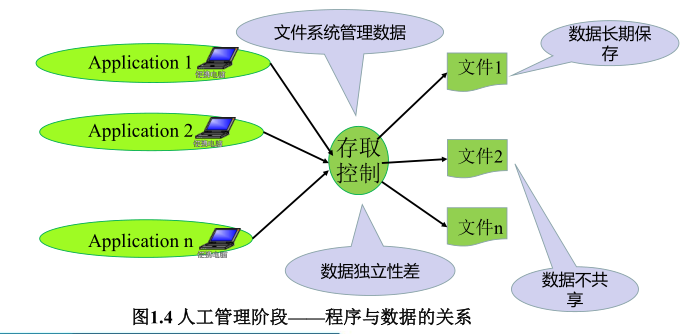
数据的共享程度：无共享、冗余度极大

数据的独立性：不独立，完全依赖于程序

数据的结构化：无结构

数据控制能力：应用程序自己控制

1. **文件系统阶段**



特点：数据的管理者：文件系统，数据可长期保存

数据面向的对象：某一应用程序

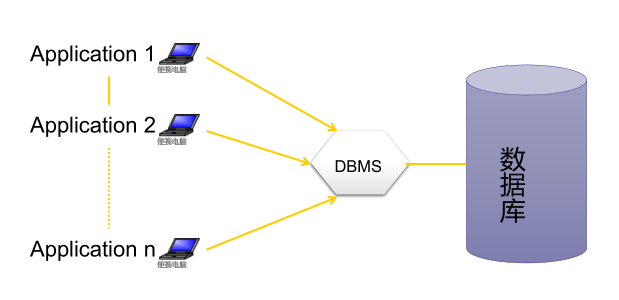
数据的共享程度：共享性差、冗余度大

数据的结构化：记录内有结构,整体无结构

数据的独立性：独立性差，数据的逻辑结构改变必须 修改应用程序

数据控制能力：应用程序自己控制

1. **数据库系统阶段**



特点：

1. 数据结构化

整体数据的结构化是数据库的主要特征之一

整体机构化：不再仅仅针对某一个应用，而是面向全组织；不仅数据内部结构化，整体是结构化的，数据之间具有联系；数据可以变长；数据的最小存取单位是数据项

数据的结构用数据模型描述，无需程序定义和解释

1. 数据的共享性高，冗余度低，易扩充

数据库系统从整体角度看待和描述数据，数据面向整个系统，可以被多

个用户、多个应用共享使用

数据共享的好处：减少数据冗余，节约存储空间；避免数据之间的不相容性与不一致性；使系统易于扩充

1. 数据独立性高

物理独立性：指用户的应用程序与存储在磁盘上的数据库中数据是相互独立的。当数据的物理存储改变了，应用程序不用改变

逻辑独立性：指用户的应用程序与数据库的逻辑结构是相互独立的。数据的逻辑结构改变了，用户程序也可以不变。

数据独立性是由DBMS的二级映像功能来保证的

1. 数据由DBMS统一管理和控制

DBMS提供的数据控制功能：

数据的安全性（Security）保护：保护数据，以防止不合法的使用造成的数据的泄密和破坏

数据的完整性（Integrity）检查：将数据控制在有效的范围内，或保证数据之间满足一定的关系

并发（Concurrency）控制：对多用户的并发操作加以控制和协调，防止相互干扰而得到错误的结果

数据库恢复（Recovery）：将数据库从错误状态恢复到某一已知的正确状态

1. 数据模型
2. 概念

数据模型是对现实世界数据特征的抽象（现实世界的模拟）

能比较真实地模拟现实世界

数据模型应满足的三方面要求 容易为人所理解

便于在计算机上实现

数据模型是数据库系统的核心和基础

1. 两大类模型

根据模型应用目的不同，**数据模型**分为两类（分属两个不同的层次）

1. 概念模型，也称信息模型，它是从用户的角度来对数据和信息建模，用于

数据库设计

1. 逻辑模型和物理模型

逻辑模型主要包括网状模型、层次模型、关系模型、面向对象模型等，按计算机系统的观点对数据建模，用于DBMS实现

物理模型是对数据最底层的抽象，描述数据在系统内部的表示方式和存取方法，在磁盘或磁带上的存储方式和存取方法

1. 两步抽象

现实世界——概念模型——物理模型

1. 数据模型组成要素

数据结构

数据操作

数据的约束条件

1. **数据结构**

数据结构：描述数据库的组成对象，以及对象之间的联系

数据结构是对系统静态特性的描述

根据数据结构的类型来命名数据模型：

非关系模型 层次模型、网状模型

关系模型 数据结构：关系

面向对象模型 数据结构：对象

1. **数据操作**

数据操作：对数据库中各种对象(型)的实例(值)允许执行的操作及有关的操作规则

增删改查

数据操作是对系统动态特性的描述

1. **数据的完整性约束条件**

一组完整性规则的集合

完整性规则：给定的数据模型中数据及其联系所具有的制约和储存规则

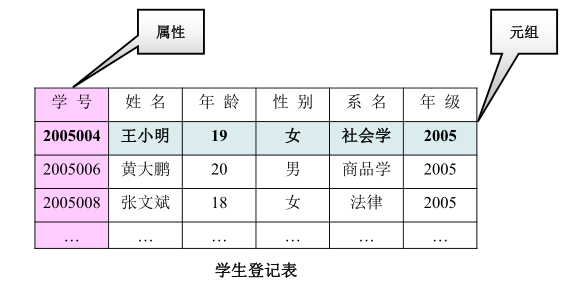
1. 关系模型

关系数据库系统采用关系模型作为数据的组织方式

1. 关系数据模型的数据结构

在用户观点下，关系模型中数据的逻辑结构是一张二维表，它由行和列

组成



关系：一个关系对应通常说的一张表

元组：表中的一行即为一个元组（属性有名字，元组无名字）

属性：表中的一列即为一个属性，给每一个属性起一个名称即属性名

码：唯一确定一个元组的属性或属性组（学号+课程号才能确定一个元组）

域：是一组具有相同数据类型的值的集合,属性的取值范围来自某个域。

分量：元组中的一个属性值

关系模式：对关系的描述，一般表示为

关系名（属性1，属性2，……，属性n）

学生（学号，姓名，年龄，性别，系，年级）

关系必须是规范化的，满足一定的规范条件。最基本的规范条件：关系的每一个分量必须是一个不可分的数据项（即表中不允许还有表）

所有的列/行都要各不相同，即不能冗余。列/行顺序可变

1. 关系数据模型的操纵和完整性约束

查询、插入、删除、更新

数据操作是集合操作，操作对象和操作结果都是关系，即若干元组的集合

存取路径对用户隐蔽，用户只要指出“干什么”，不必详细说明“怎么干”

实体完整性

关系的完整性约束条件 参照完整性

用户定义完整性

1. 数据库系统结构
2. 概念
3. 从数据库应用开发人员角度看，数据库系统通常采用三级模式结构，是数据库系统内部的系统结构
4. “型” 和“值” 的概念

型：对某一类数据的结构和属性的说明

值：是型的一个具体赋值

例如：

学生记录型：（学号，姓名，性别，系别，年龄，籍贯）

一个记录值：（201315130，李明，男，计算机，19，江苏）

1. 模式：数据库逻辑结构和特征的描述、是型的描述、反映的是数据的结构及其联系、模式是相对稳定的

实例：模式的一个具体值、反映数据库某一时刻的状态、同一个模式可以有很多实例、实例随数据库中的数据的更新而变动

例如：在学生选课数据库模式中，包含学生记录、课程记录和学生选课记录.

2013年的一个学生数据库实例，包含：

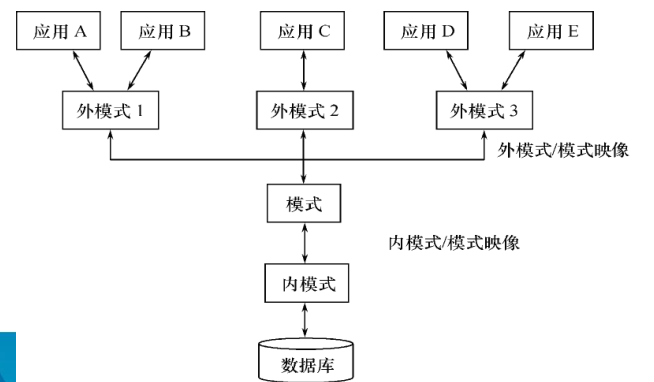
2013年学校中所有学生的记录

学校开设的所有课程的记录

所有学生选课的记录

2012年度学生数据库模式对应的实例与 2013年度学生数据库模式对应的实例是不同的

1. 数据库系统的三级模式结构



1. **模式（逻辑模式）**

数据库中全体数据的**逻辑结构和特征**的描述

所有用户的公共数据视图，综合了所有用户的需求

**一个数据库只有一个模式**

模式的地位：是数据库系统模式结构的中间层，独立于数据库的其他层次，与数据的物理存储细节和硬件环境无关，与具体的应用程序、开发工具及高级程序设计语言无关

数据的逻辑结构（数据项的名字、类型、取值范围等）、数据之间的联系、数据有关的安全性、完整性要求

全局逻辑结构是数据库的中心与关键，设计数据库模式结构时应首先确定数据库的逻辑模式

1. **外模式（也称子模式或用户模式）**

数据库用户（包括应用程序员和最终用户）使用的局部数据的逻辑结构和特征的描述，**数据库用户的数据视图，是与某一应用有关的数据的逻辑表示**

外模式地位：介于模式与应用之间

模式与外模式的关系：一对多

外模式通常是模式的子集

一个数据库可以有多个外模式。反映了不同的用户的应用需求、看待数据的方式、对数据保密的要求

对模式中同一数据，在外模式中的结构、类型、长度、保密级别等都可以不同

外模式与应用的关系：一对多

同一外模式也可以为某一用户的多个应用系统所使用

但一个应用程序只能使用一个外模式

外模式的用途：保证数据库安全性的一个有力措施，每个用户只能看见和访问所对应的外模式中的数据

面向具体的应用程序、定义在逻辑模式之上、独立于存储模式和存储设备、当应用需求发生较大变化，相应外模式不能满足其视图要求时，该外模式就得做相应改动、设计外模式时应充分考虑到应用的扩充性

1. **内模式（也称存储模式）**

是数据物理结构和存储方式的描述，是数据在数据库内部的表示方式

记录的存储方式（顺序存储，按照B树结构存储，按hash方法存储）、索引的组织方式、数据是否压缩存储、数据是否加密、数据存储记录结构的规定

**一个数据库只有一个内模式**

依赖于它的全局逻辑结构、独立于数据库的用户视图，即外模式、独立于具体的存储设备、将全局逻辑结构中所定义的数据结构及其联系按照一定的物理存储策略进行组织，以达到较好的时间与空间效率

1. 特定的应用程序

在外模式描述的数据结构上编制的、依赖于特定的外模式、与数据库的模式和存储结构独立、不同的应用程序有时可以共用同一个外模式

1. 数据库的二级映像功能与数据独立性
2. 三级模式是对数据的三个抽象级别

外模式：描述的是数据的**局部逻辑结构**

模式：描述的是数据的**全局逻辑结构**

内模式：数据在数据库内部的存储方式

1. 二级映象在数据库管理系统内部实现这三个抽象层次的联系和转换

外模式／模式映像

模式／内模式映像

保证了数据库外模式的稳定性、从底层保证了应用程序的稳定性，除非应用需求本身发生变化，否则应用程序一般不需要修改

1. 外模式／模式映象

定义外模式与模式之间的对应关系

**同一个模式可以有任意多个外模式**

外模式／模式映象通常包含在**各自外模式**的描述中

**保证数据的逻辑独立性：**当模式改变时，数据库管理员修改有关的外模式／模式映象，使外模式保持不变、应用程序是依据数据的外模式编写的，从而应用程序不必修改，保证了数据与程序的逻辑独立性，简称数据的逻辑独立性。

1. 模式／内模式映象

定义了数据全局逻辑结构与存储结构之间的对应关系。例如:说明逻辑记录和字段在内部是如何表示的

**数据库中模式／内模式映象是唯一的**

该映象定义通常包含在**模式**描述中

**保证数据的物理独立性：**当数据库的存储结构改变了（例如选用了另一种存储结构），数据库管理员修改模式／内模式映象，使模式保持不变、应用程序不受影响。保证了数据与程序的物理独立性，简称数据的物理独立性

数据与程序之间的独立性，使得数据的定义和描述可以从应用程序中分离出去

1. 数据库系统的组成
2. 数据库+数据库管理系统（及其开发工具）+应用程序+数据库管理员+硬件平台及数据库+软件+人员
3. 