



2024年春季学期

数据库系统及应用

陈欢欢，周熙人



本课程研究问题

教务信息
系统



证券信息
系统



银行信息
系统

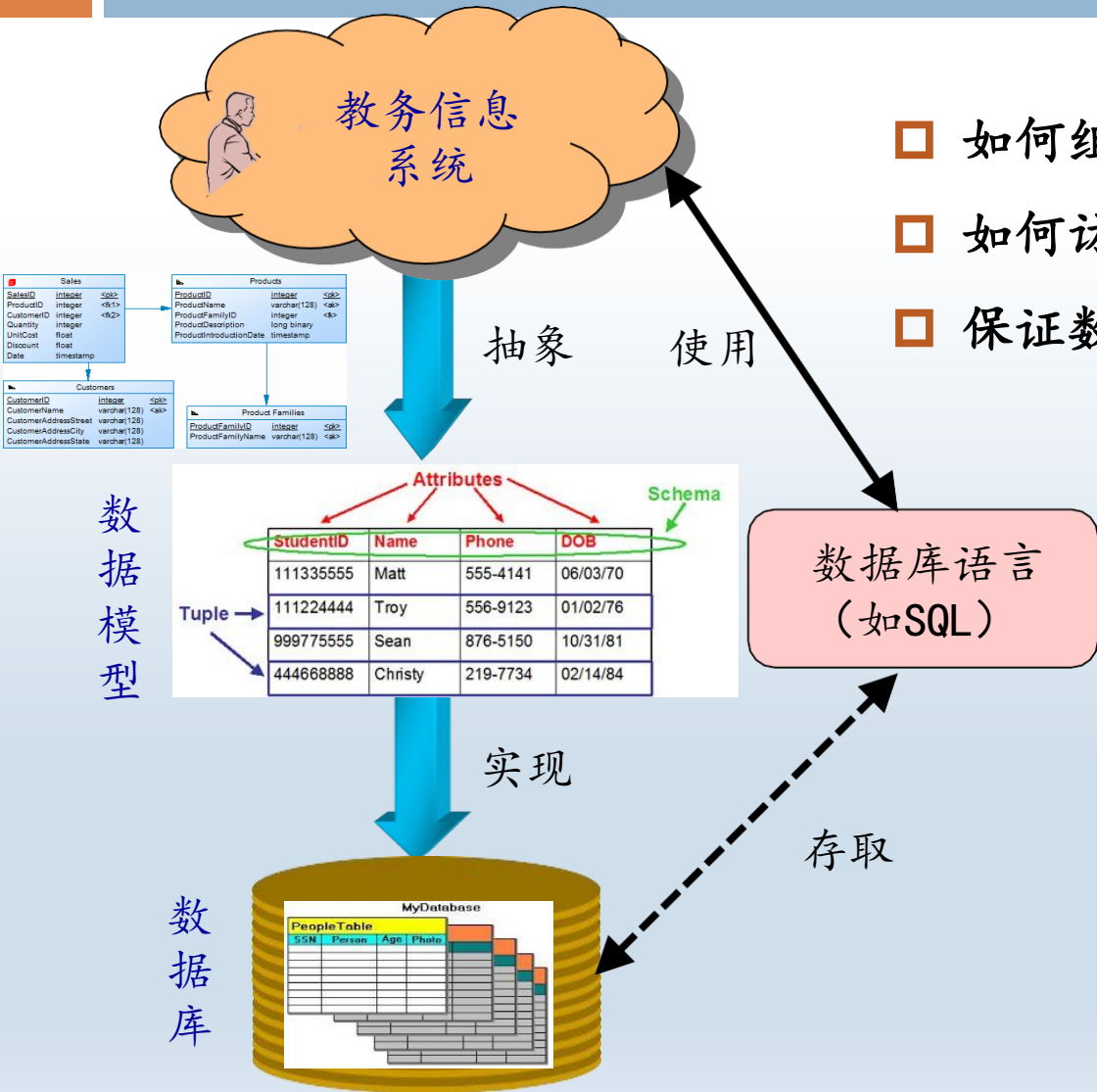


数据库

- 数据量大
- 数据联系复杂
- 数据共享程度高
- 存取性能要求高
- 数据一致性要求高
- 数据类型多样化



本课程研究的问题



- 如何组织并存储到数据库中？
- 如何访问数据库中的数据？
- 保证数据的一致性、安全性和效率？



课程主要内容

□ 数据如何组织并存储到数据库中

- 数据库体系结构
- 数据模型
- 数据库设计



数据组织

□ 应用如何访问数据库中的数据？

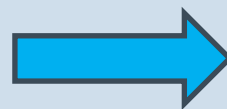
- SQL和PL/SQL
- 数据库应用编程



数据存取

□ 如何保证数据的一致性、安全性和存取效率？

- 事务管理
- 恢复、并发控制、完整性、安全性
- 索引、查询优化等



数据管理



课程知识结构

Chp. 1 数据库系统概述

Chp. 2 数据库系统体系结构

Chp. 3 关系数据模型

Chp. 10 事务与恢复

Chp. 4 SQL

Chp. 6 关系数据库模式设计

Chp. 11 并发控制

Chp. 5 过程化SQL

Chp. 7 数据库设计

Chp. 12 数据库安全性

Chp. 8 数据库索引

Chp. 13 数据库完整性

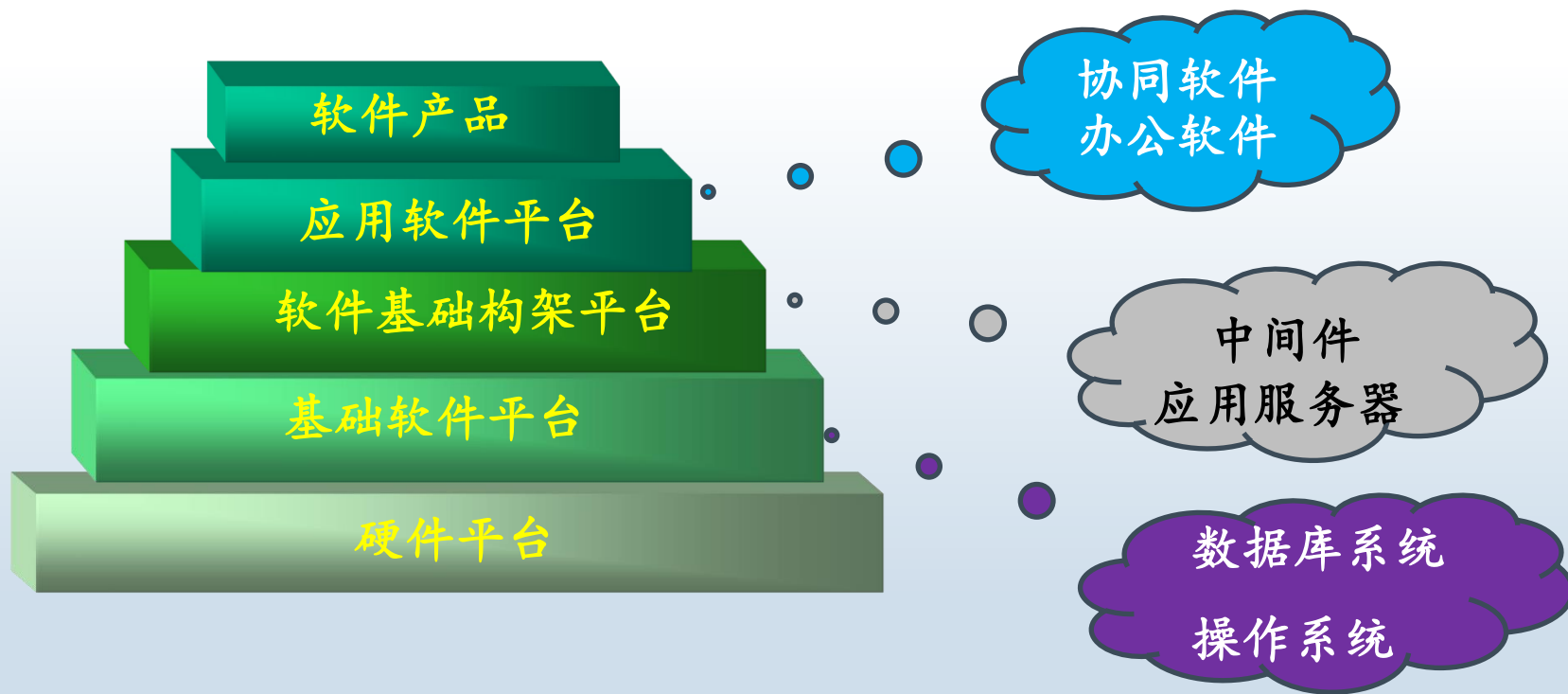
* Chp. 9 数据库应用开发



Chp. 14 高级主题



“数据库”在计算机系统的位置





课程目的

- 基础知识方面
 - 了解数据库技术的发展概况
 - 了解数据库系统的相关概念
- 数据模型和数据库设计方面
 - 掌握关系数据库理论
 - 掌握关系数据库的模式设计方法
 - 熟练掌握关系数据库的规范化设计方法
- 数据存取方面
 - 熟练掌握数据库查询语言SQL、过程化SQL、数据库应用开发
- 数据库系统方面
 - 基本掌握数据库保护技术：恢复、并发控制、完整性、安全性
- 了解NoSQL等高级数据库技术



与高阶课程之间的关系

本科课程

数据库原理/
数据库系统及应用

Application-Oriented

硕士课程

高级数据库系统

System-Oriented





课程重点和难点

□ 理解关系数据模型的设计原理

□ 凭什么获图灵奖？[1981，图灵奖]

□ 如何用关系代数和SQL正确表达用户查询？

cnki 中国知网 cnki.net

文献 期刊 博硕士 会议 报纸 外文文献 年鉴 百科 词典 统计数据 专利 标准 更多>>

高级检索 专业检索 作者发文检索 句子检索 一框式检索

新型出

»文献分类目录

全选 清除

- ☒ 基础科学
- ☒ 工程科技 I 辑
- ☒ 工程科技 II 辑
- ☒ 农业科技
- ☒ 医药卫生科技
- ☒ 哲学与人文科学
- ☒ 社会科学 I 辑
- ☒ 社会科学 II 辑
- ☒ 信息科技

输入检索条件：

☐ ☐ (主题 数据库 词频 并含 词频 精确)

并且 (篇名 词频 并含 词频 精确)

☐ ☐ 作者 中文名/英文名/拼音 精确 作者单位： 全称/简称/曾用名 模糊

发表时间： 从 到 更新时间： 不限

文献来源： 模糊

支持基金： 模糊

☐ 网络首发 ☐ 增强出版 ☐ 数据论文 ☐ 中英文扩展 ☐ 同义词扩展

检索



课程重点和难点

- 如何评价数据库设计的好坏？如何给出一个好的数据库设计？
- 理解DBMS的事务处理机制 [1998, 图灵奖]



2024年春季学期

第1章 数据库系统概述



主要内容

- 数据库系统的基本概念
- 为什么使用数据库?
- DBMS的功能
- DBMS的分类
- DBMS的架构
- 数据库技术的发展



一、数据库系统的基本概念

- 数据
- 数据库
- 数据库模式
- 数据库管理系统
- 数据库系统



1、数据

- 数据(Data)是数据库中存储的基本对象
- 数据的定义
 - 人们用来反映客观世界而记录下来的可以鉴别的符号
- 数据的种类
 - 数值数据：0—9
 - 非数值数据：字符、文字、声音、图形、图像等





1、数据

- 数据的特点
 - 数据与其语义是不可分的
- 例子1：93是一个数据
 - 语义1：学生某门课的成绩
 - 语义2：某人的体重
 - 语义3：2016级学生人数
- 例子2：学生档案记录（李明，199705，中国科大，2010）
 - 语义1：学生姓名，出生年月，所在学校，毕业年份
 - 语义2：学生姓名，出生年月，录取大学，入学时间



学生档案中的学生记录:

(李明, 男, 199505, 江苏南京市, 计算机系, 2013)

语义: 学生姓名、性别、出生年月、出生地、所在院系、入学时间

解释: 李明是个大学生, 1995年5月出生, 江苏南京市人, 2013年考入计算机系



2、数据库

□ 数据库的定义

□ 数据库 (Database, 简称DB) 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

□ 数据库的基本特征

□ 持久存储，一般数据量比较大

□ 数据按一定的数据模型组织、描述和储存

□ 可为各种用户共享

□ 数据间联系密切，具有最小的冗余度和较高的独立性

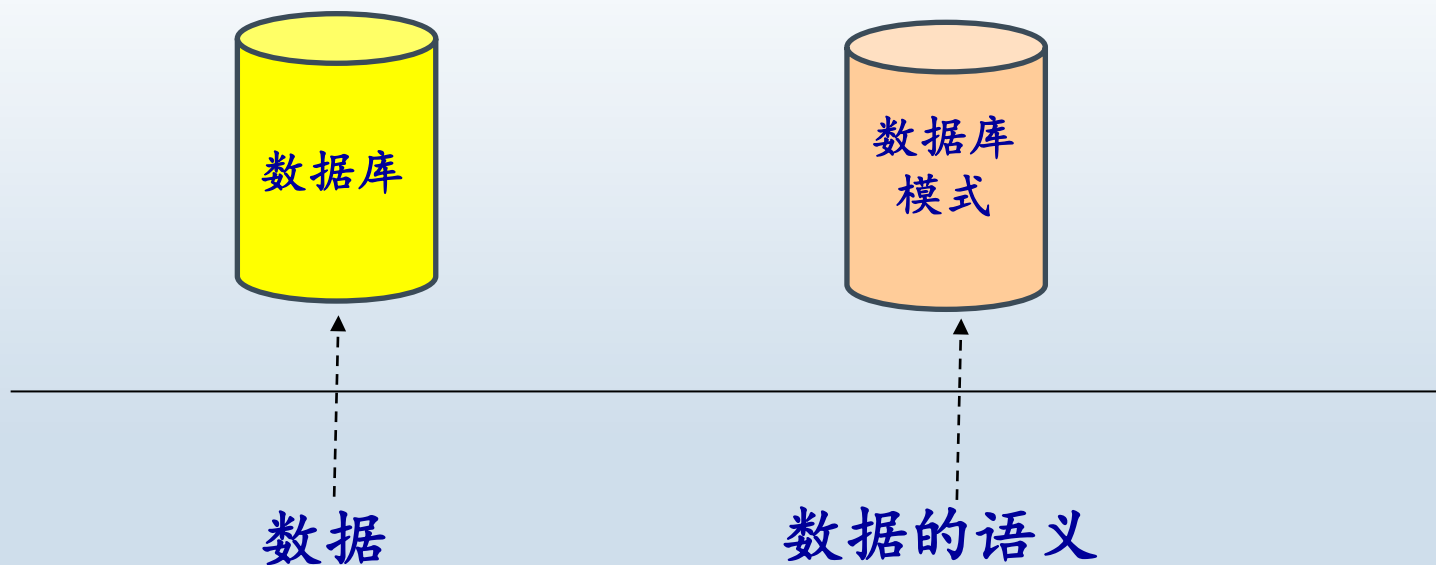
□ 服务于某个特定的应用

■ 例：图书数据库、航班数据库、银行数据库……



3、数据库模式

- 数据库模式是数据库中全体数据的逻辑结构和特征的描述





举例

学生(学号:char, 姓名:char, 年龄:int)

学号	姓名	年龄
001	张三	20
002	李四	21
003	王五	22



学号	姓名	年龄
----	----	----

模式

001	张三	20
002	李四	21
003	王五	22

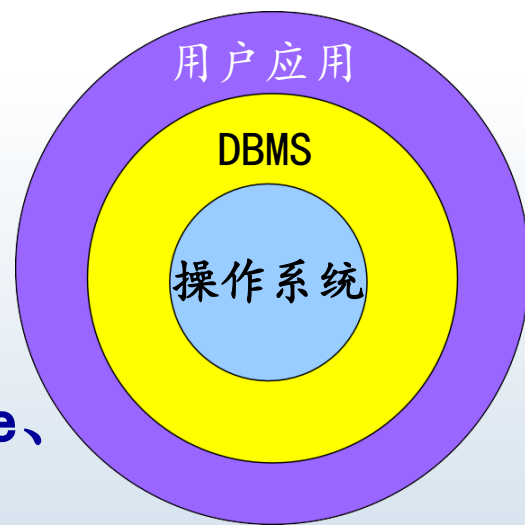
数据



4、数据库管理系统

□ **DBMS (Database Management System)**，是计算机程序的集合，用于创建和维护数据库

- 位于操作系统和用户应用之间
- 总是基于某种数据模型
- 数据库厂商的产品通常指DBMS，如Oracle、SQL Server、MySQL、DB2、Informix、PostgreSQL等



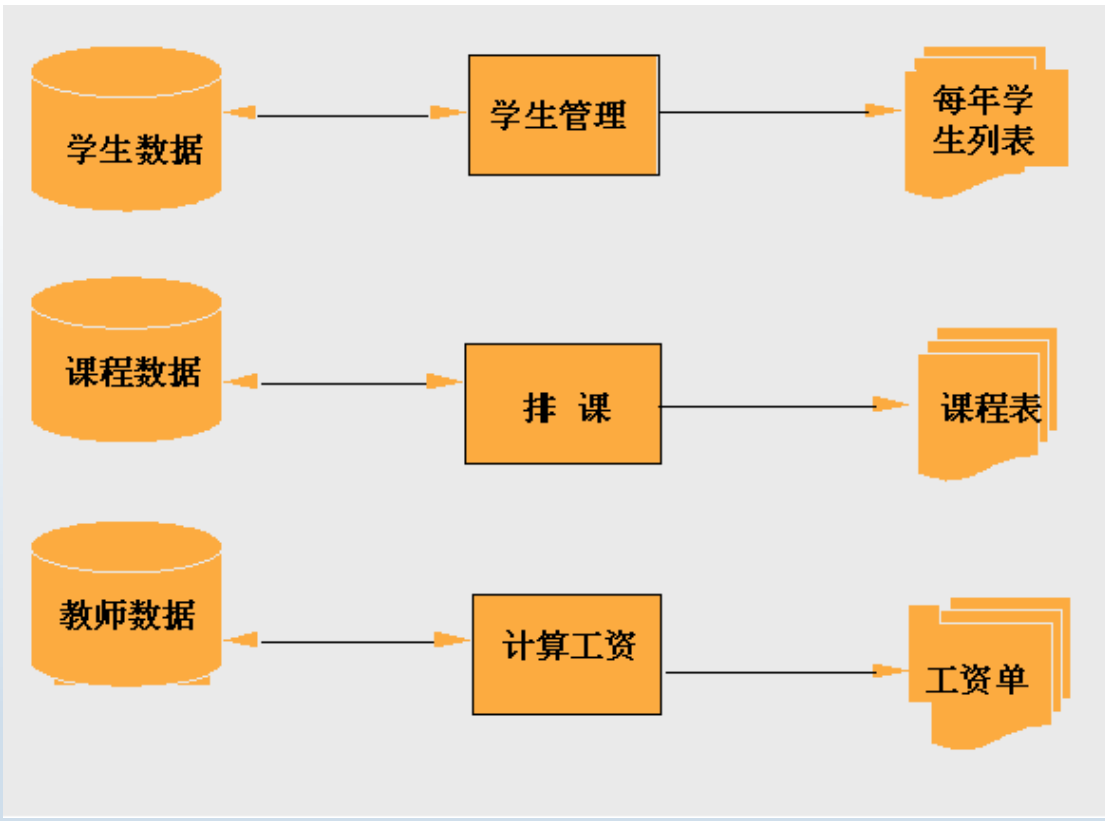


5、数据库系统

- DBS (DataBase System)，指在计算机系统中引入了数据库后的系统，即采用了数据库技术的计算机系统
 - 数据：数据库中的数据
 - 硬件：二级存储器、处理器、主存等计算机硬件
 - 软件：DBMS、应用系统
 - 用户
 - 应用程序员：使用C++、Java等程序设计语言编写数据库应用程序
 - 终端用户：通过联机工作站或终端与数据库系统交互，一般使用特定的语言和界面
 - 数据库管理员 (DBA)：负责对系统资源的管理和维护
 - 数据库设计员和系统分析员



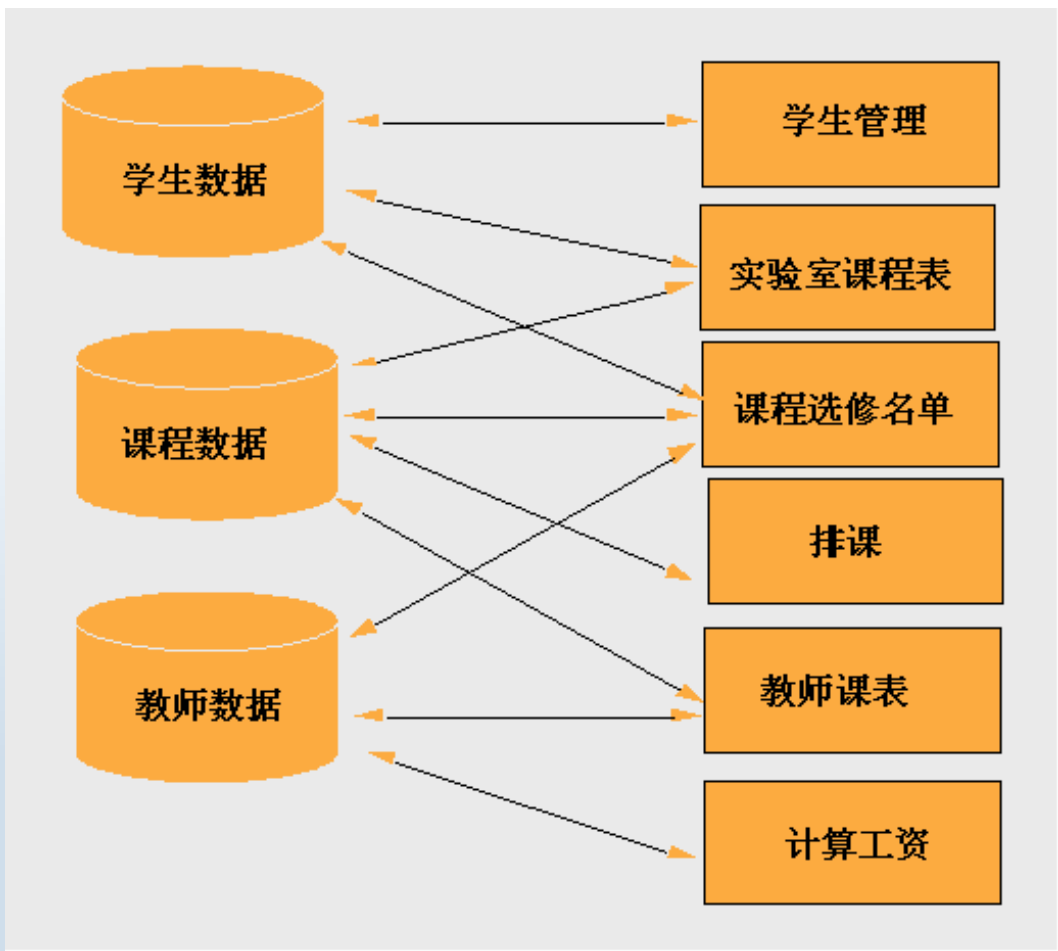
二、为什么使用数据库？



无共享数据时的文件系统



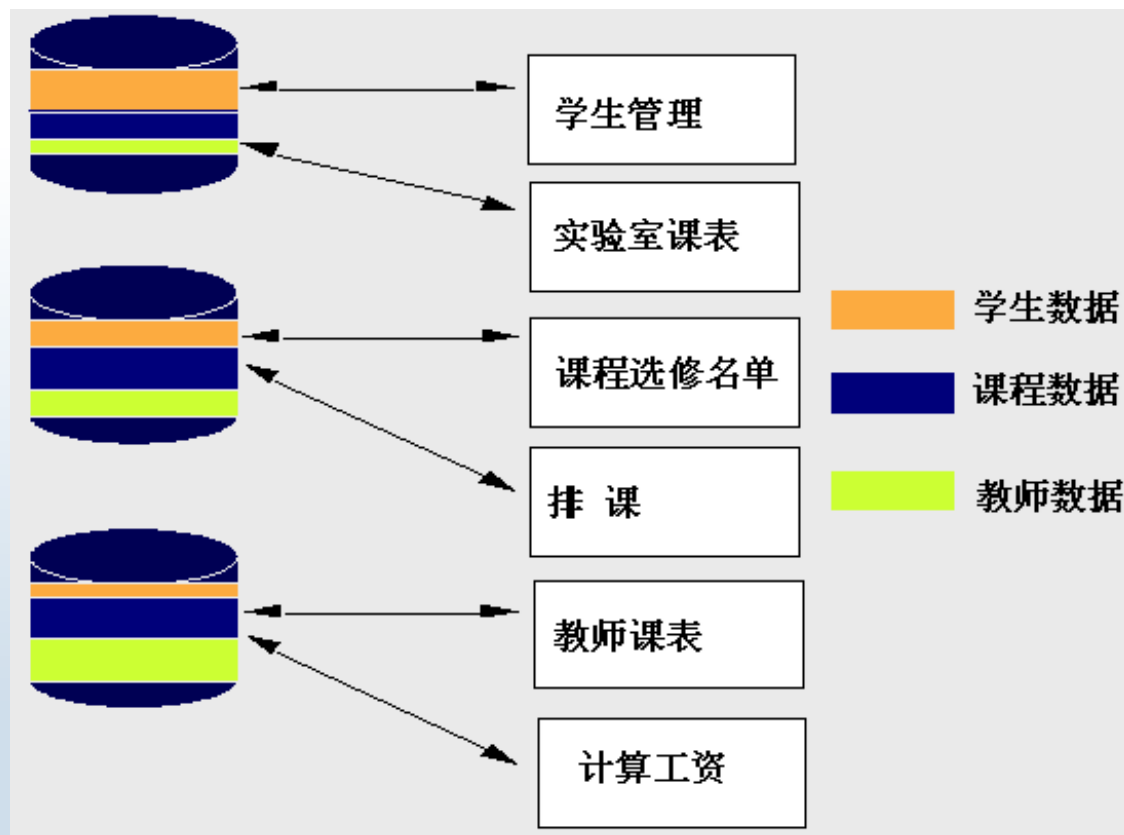
二、为什么使用数据库？



有共享数据时的文件系统



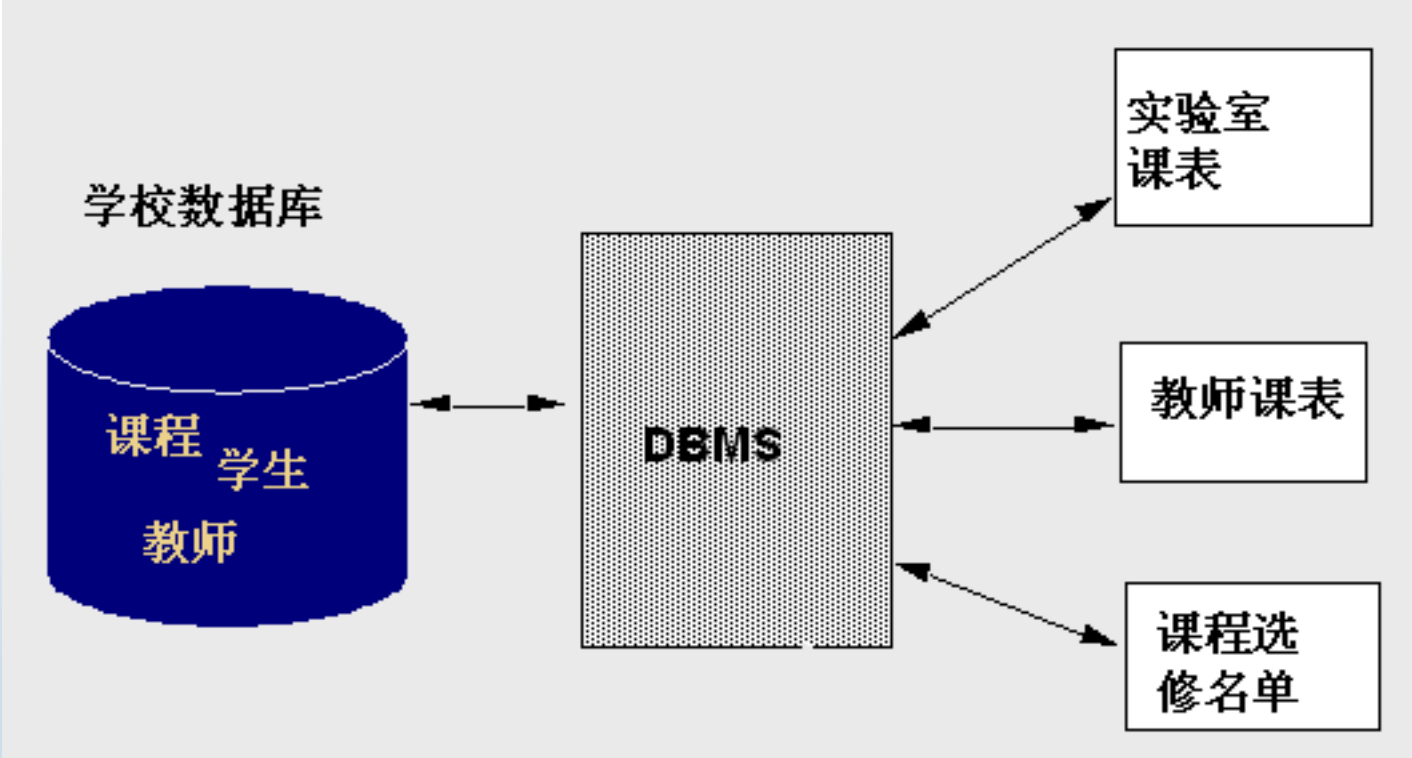
二、为什么使用数据库？



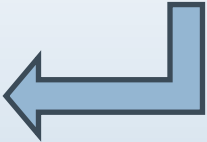
- 数据冗余和不一致
- 数据访问困难
- 数据孤立
- 完整性问题
- 原子性问题
- 并发访问异常
- 安全性问题



二、为什么使用数据库？



使用数据库减少冗余，避免不一致





二、为什么使用数据库？

- 数据共享
- 减少冗余
- 避免不一致
- 提供事务支持：如银行转帐
 - 原子性、一致性、隔离性、持久性
- 保持完整性
- 增强安全性
- 提供并发控制
- 标准化



三、DBMS的功能

- 数据库定义
 - 数据库对象定义：表、索引、约束、用户等
- 数据库操纵
 - 实现对数据库的基本操作：增、删、改、查
- 数据库保护
 - 恢复、并发控制、完整性控制、安全性控制
- 数据库的建立和维护
 - 初始数据的转换和装入、数据备份、数据库的重组织、性能监控和分析等
 - 通常由一些实用程序完成



四、DBMS的分类

□ 按数据模型

- 网状型DBMS
 - 层次型DBMS
 - 关系型DBMS
 - 对象DBMS
 - NoSQL
- 第1代DBMS
- 第2代DBMS
- 第3代DBMS* (非公认)
- 第4代DBMS?

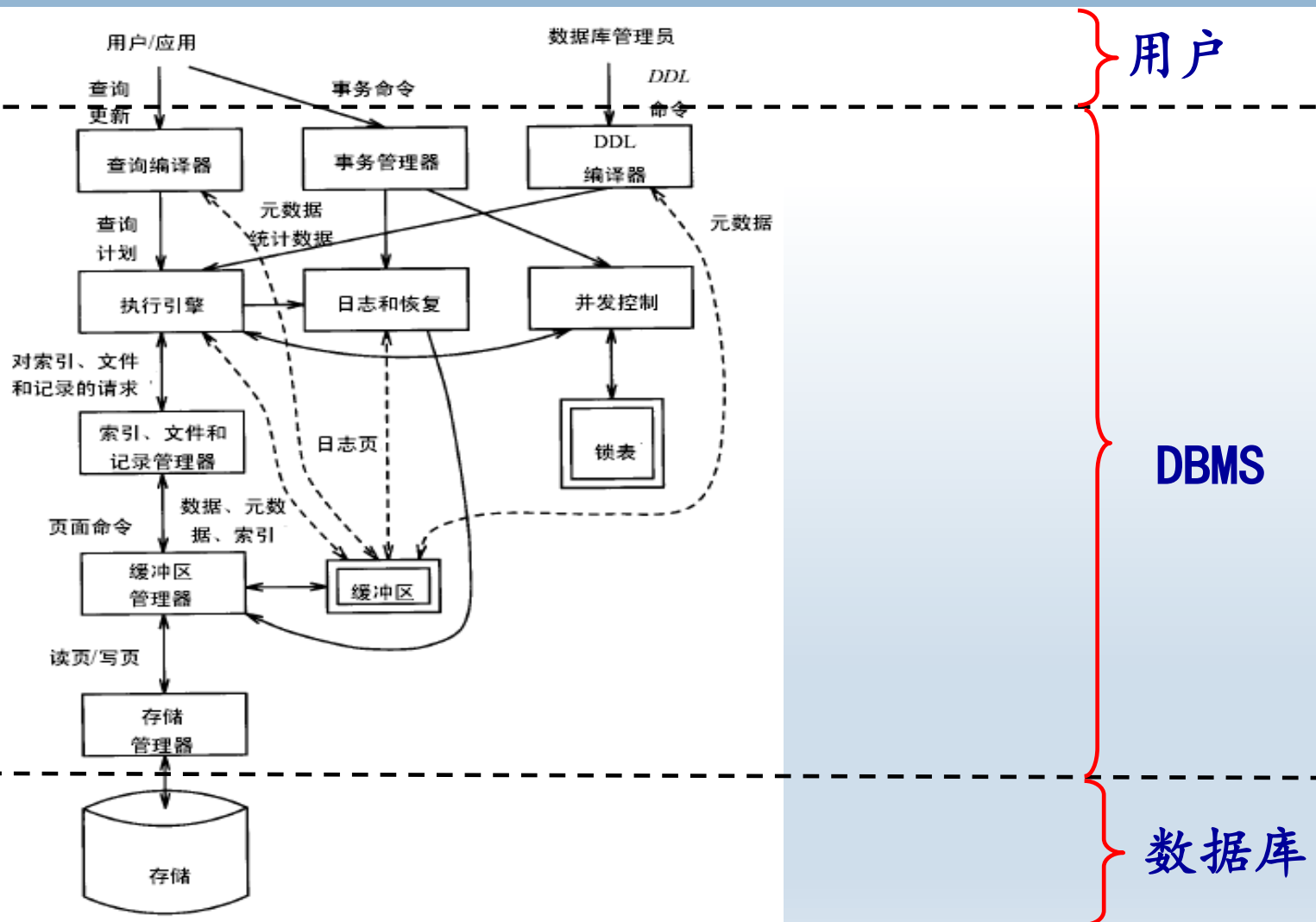


四、DBMS的分类

- 按所支持的用户数
 - 单用户DBMS（目前已经很少见）
 - 多用户DBMS
- 按允许数据库可以分布的站点数
 - 集中式DBMS
 - 分布式DBMS
- 按用途
 - 通用DBMS，如Oracle、Informix等
 - 专用DBMS，如时态数据库、空间数据库、移动数据库等



五、DBMS的架构





六、数据管理技术的发展

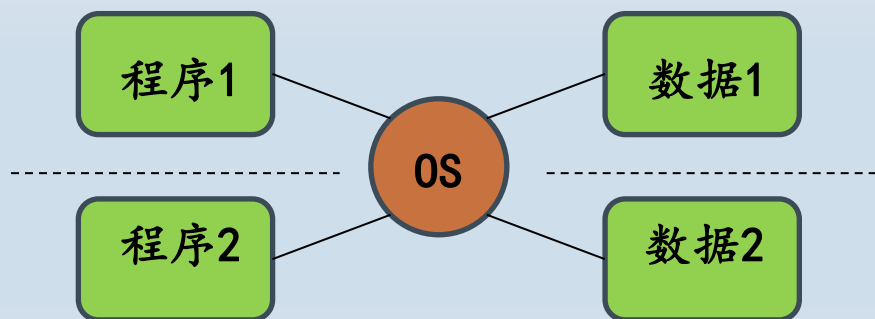
- 人工管理阶段（20世纪50年代中以前）
 - 数据不保存在机器中
 - 应用程序自己管理数据
 - 数据无共享
 - 只有程序概念，没有文件概念





六、数据管理技术的发展

- 文件系统阶段（20世纪50s后—60s中）
 - 数据可以长期保存在磁盘上
 - 文件系统管理数据
 - 数据共享性差，冗余大：冗余时必须建立不同的文件以满足不同的应用
 - 数据独立性差：程序通过文件名即可访问数据，但文件结构改变时必须修改程序



文件系统阶段应用程序与数据之间的对应关系



文件系统阶段（补充）

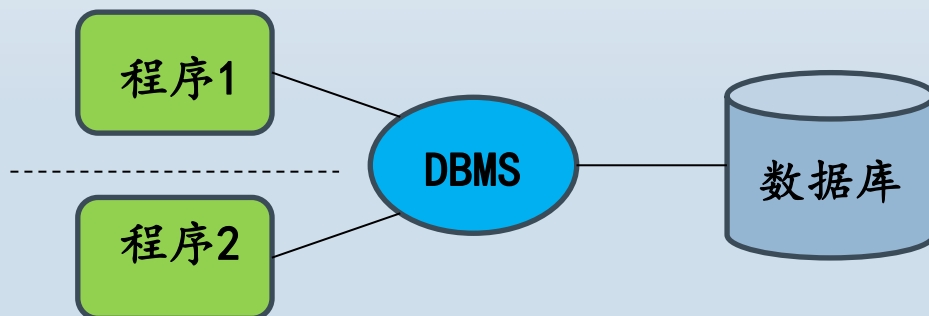
□ 特点

- 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
- 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
- 数据控制能力：应用程序自己控制



六、数据管理技术的发展

- 数据库系统阶段（20世纪60s末——）
 - 数据结构化：采用复杂数据模型，不仅可以表示数据，还可以表示数据间的联系
 - 高共享，低冗余
 - 数据独立性高
 - 数据由DBMS统一控制





六、数据管理技术的发展（补充）

- 时期
 - 20世纪60年代末以来
- 产生的背景
 - 应用背景：大规模数据管理
 - 硬件背景：大容量磁盘、磁盘阵列
 - 软件背景：有数据库管理系统
 - 处理方式：联机实时处理, 分布处理, 批处理
- 文件系统到数据库系统，标志着数据管理技术的飞跃



七、数据库技术的发展

- 1961: GE的C. W. Bachman设计了历史上第一个DBMS——网状数据库系统IDS (Integrated DataStore) [1973, 图灵奖]
- 1968: IBM设计了层次数据库系统IMS
- 1969: CODASYL的DBTG发表了网状数据模型报告, 奠定了网状数据库技术
- 1970: IBM的E. F. Codd提出了关系数据模型, 奠定了关系数据库理论基础 [1981, 图灵奖]
- 1974: IBM的Boyce和Chamberlin设计了SQL语言
- 1973~1976: E. F. Codd设计了System R, M. Stonebraker设计了Ingres
- 1976: IBM的Jim Gray提出了一致性、锁粒度等设计, 奠定了事务处理基础 [1998, 图灵奖]
- 1977: Larry Ellison创建了Oracle公司, 1979年发布Oracle 2.0, 1986年Oracle上市
- 1983: IBM发布DB2



七、数据库技术的发展

- 1985: 面向对象数据库技术提出
- 1987: Sybase 1.0发布
- 1990: M. Stonebraker发表“第三代数据库系统宣言”，提出对象关系数据模型 [2014, 图灵奖] “For fundamental contributions to the concepts and practices of underlying modern database systems”
- 1987~1994: Sybase和Microsoft合作，发布 Sybase SQL Server 4.2。破裂后Sybase继续发布Sybase ASE 11.0
- 1996: Microsoft发布Microsoft SQL Server 6.5
- 1996: 开源的MySQL正式发布
- 1998: 提出了半结构化数据模型 (XML1.0)
- 2005, M. Stonebraker等开发完成C-Store, Column-based DBMS
- 2007, NoSQL (非关系型数据库) 在Web领域大行其道。
Amazon (SimpleDB/Dynamo), Google (BigTable/LevelDB),
Facebook (RocksDB/Cassandra), MongoDB, HBase, Redis, etc.



八、数据库领域的出版物

□ 国际会议

- A类：SIGMOD、VLDB、ICDE (DB三大会议)
- B类：EDBT、ICDT、CIDR、CIKM、DASFAA
- C类：DEXA、APWeb-WAIM、ER、SSTD、SSDBM、MDM、WebDB、ADBIS等
- 中国数据库学术会议NDBC

□ 国际期刊

- A类：VLDB Journal、TKDE、TODS (DB三大期刊)
- B类：DKE、Information Systems、GeoInformatica等

可参考中科院分区表和CCF计算机国际会议与期刊排名



本章小结

- ❑ 数据库系统的基本概念
- ❑ 文件系统和数据库系统
- ❑ DBMS的功能
- ❑ DBMS的分类
- ❑ DBMS的架构
- ❑ 数据库技术的发展