

数据库系统概论

An Introduction to Database Systems

第一章 绪论

重点：

牢固掌握概念模型的基本概念及其主要建模方法 ER 方法；

掌握关系数据模型的相关概念、数据库系统三级模式和两层映像的体系结构；

数据库系统的逻辑独立性和物理独立性等；

难点：

本章的难点是需要掌握数据库领域大量的基本概念。

数据模型及数据库系统的体系结构也是本章的难点。

内容概述

阐述数据库的基本概念，介绍数据管理技术的进展情况、数据库技术产生和发展的背景，数据模型的基本概念、组成要素和主要的数据库模型，概念模型的基本概念及 ER 方法，数据库系统的 3 级模式结构以及数据库系统的组成。

本章目标

本章讲解的数据库基本概念和基本知识

学习本章的目的在于了解基本知识，掌握基本概念，为以后的学习打好扎实的基础。

第一章 绪论

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统结构

1.4 数据库系统的组成

1.5 小结

1.1 数据库系统概述

1.1.1 数据库的地位

1.1.2 四个基本概念

1.1.3 数据管理技术的产生与发展

1.1 数据库系统概述

1.1.1 数据库的地位

1.1.2 四个基本概念

1.1.3 数据管理技术的产生与发展

数据库的地位

- 数据库技术产生于六十年代末，是数据管理的最新技术，是计算机科学的重要分支。
- 数据库技术是信息系统的核心和基础，它的出现极大地促进了计算机应用向各行各业的渗透。
- 数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

1.1 数据库系统概述

1.1.1 数据库的地位

1.1.2 四个基本概念

1.1.3 数据管理技术的产生与发展

1.1.2 四个基本概念

☞ 数据

☞ 数据库

☞ 数据库管理系统

☞ 数据库系统

一、数据

☞ 数据 (Data) 是数据库中存储的基本对象

☞ 数据的定义

— 描述事物的符号记录

☞ 组成数据的符号种类

— 数字、字符串、日期、逻辑值、文本、图形、图象、声音

☞ 数据的特点

— 数据与其语义是不可分的

数据举例

☞ 学生档案中的学生记录

(李明, 男, 1972, 江苏, 计算机系, 1990)

☞ 数据的形式不能完全表达其内容

☞ 数据的解释

- 语义：学生姓名、性别、出生年份、籍贯、所在系别、入学时间
- 解释：李明是个大学生，1972年出生，男，江苏人，1990年考入计算机系

二、数据库

人们收集并抽取出一个应用所需要的大量数据之后，应将其保存起来以供进一步加工处理，进一步抽取有用信息。

数据库（续）

学生登记表

学 号	姓 名	年 令	性 别	系 名	年 级
95004	王小明	19	女	社会学	95
95006	黄大鹏	20	男	商品学	95
95008	张文斌	18	女	法律学	95
...

数据库（续）

∞ 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、可共享的数据集合。

数据库（续）

数据库的特征

- 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
- 冗余度较小
- 数据独立性较高
- 易扩展

三、数据库管理系统

什么是 DBMS

- 数据库管理系统（Database Management System，简称 DBMS）是位于用户与操作系统之间的一层数据管理软件。

DBMS 的用途

- 科学地组织和存储数据、高效地获取和维护数据

数据库管理系统（续）

DBMS 的主要功能

- 数据定义功能：提供数据定义语言 (DDL)，用于定义数据库中的数据对象。
- 数据操纵功能：提供数据操纵语言 (DML)，用于操纵数据实现对数据库的基本操作（查询、插入、删除和修改）。

数据库管理系统（续）

DBMS 的主要功能（续）

- 数据库的运行管理：保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。
- 数据库的建立和维护功能：提供实用程序，完成数据库数据批量装载，数据库转储，介质故障恢复，数据库的重组和性能监视等

四、数据库系统

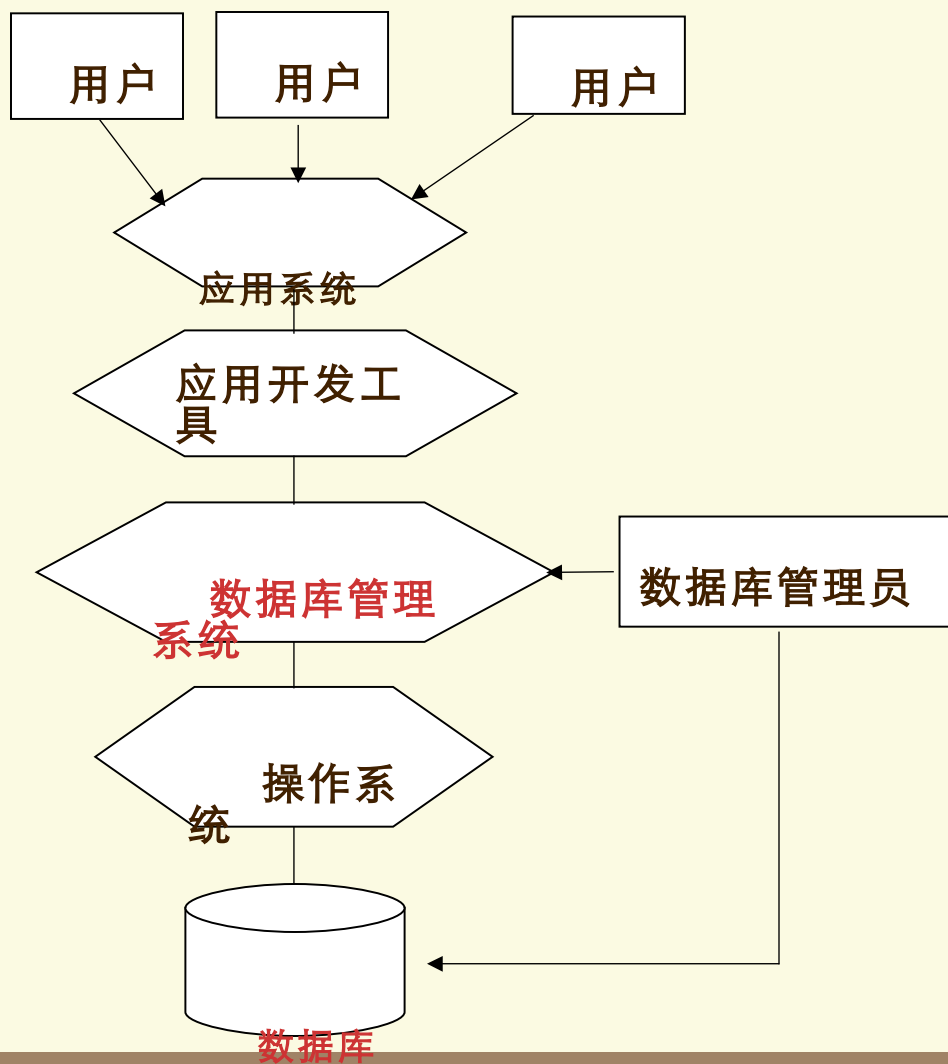
什么是数据库系统

- 数据库系统（ Database System，简称 DBS）是指在计算机系统中引入数据库后的系统构成。
- 在不引起混淆的情况下常常把数据库系统简称为数据库。

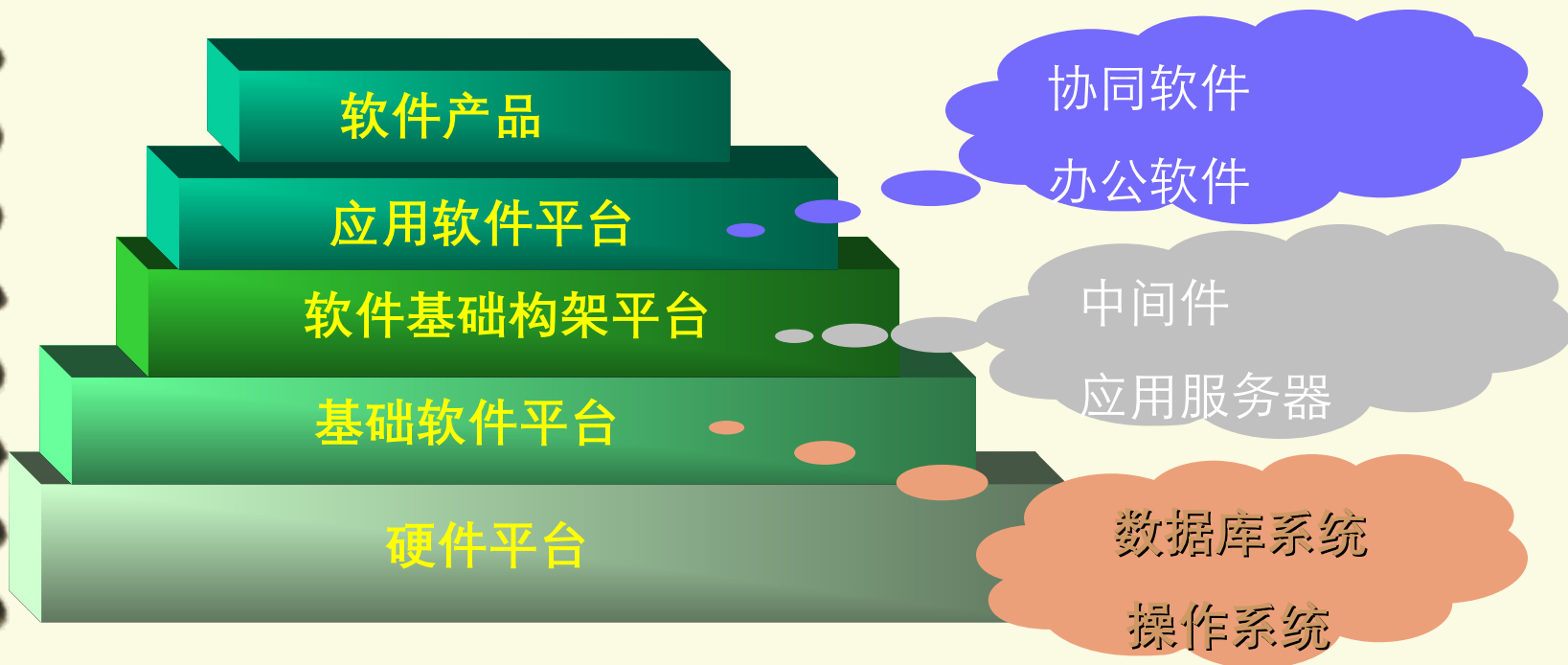
数据库系统的构成

- 由数据库、数据库管理系统（及其开发工具）、应用系统、数据库管理员构成。

数据库系统构成



数据库在计算机系统中的位置



1.1 数据库系统概述

1.1.1 数据库的地位

1.1.2 四个基本概念

1.1.3 数据管理技术的产生与发展

1.1.3 数据管理技术的产生和发展

什么是数据管理

- 对数据进行**分类**、**组织**、**编码**、**存储**、**检索**和**维护**，是数据处理的中心问题。

数据管理技术的发展过程

- 人工管理阶段 (40 年代中 --50 年代中)
- 文件系统阶段 (50 年代末 --60 年代中)
- 数据库系统阶段 (60 年代末 -- 现在)

数据管理技术的产生和发展 (续)

∞ 数据管理技术的发展动力

- 应用需求的推动
- 计算机硬件的发展
- 计算机软件的发展

		人工管理阶段	文件系统阶段	数据库系统阶段
背景	应用背景	科学计算	科学计算、管理	大规模管理
	硬件背景	无直接存取存储设备	磁盘、磁鼓	大容量磁盘
	软件背景	无操作系统	有文件系统	有数据库管理系统
	处理方式	批处理	联机实时处理，批处理	联机实时处理，分布处理，批处理
特点	数据的管理者	用户（程序员）	文件系统	数据库管理系统
	数据面向的对象	某一应用程序	某一应用	现实世界
	数据的共享程度	无共享，冗余度极大	共享性差，冗余度大	共享性高，冗余度小
	数据的独立性	不独立，完全依赖于程序	独立性差	具有高度的物理独立性和一定的逻辑独立性
	数据的结构化	无结构	记录内有结构，整体无结构	整体结构化，用数据模型描述
	数据控制能力	应用程序自己控制	应用程序自己控制	由 DBMS 提供数据安全性、完整性、并发控制和恢复能力

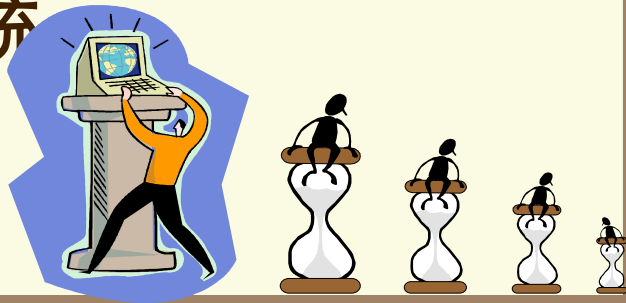
一、人工管理

时期

— 40 年代中 --50 年代中



产生的背景

- | | |
|--------|-----------|
| — 应用需求 | 科学计算 |
| — 硬件水平 | 无直接存取存储设备 |
| — 软件水平 | 没有操作系统 |
| — 处理方式 | 批处理 |

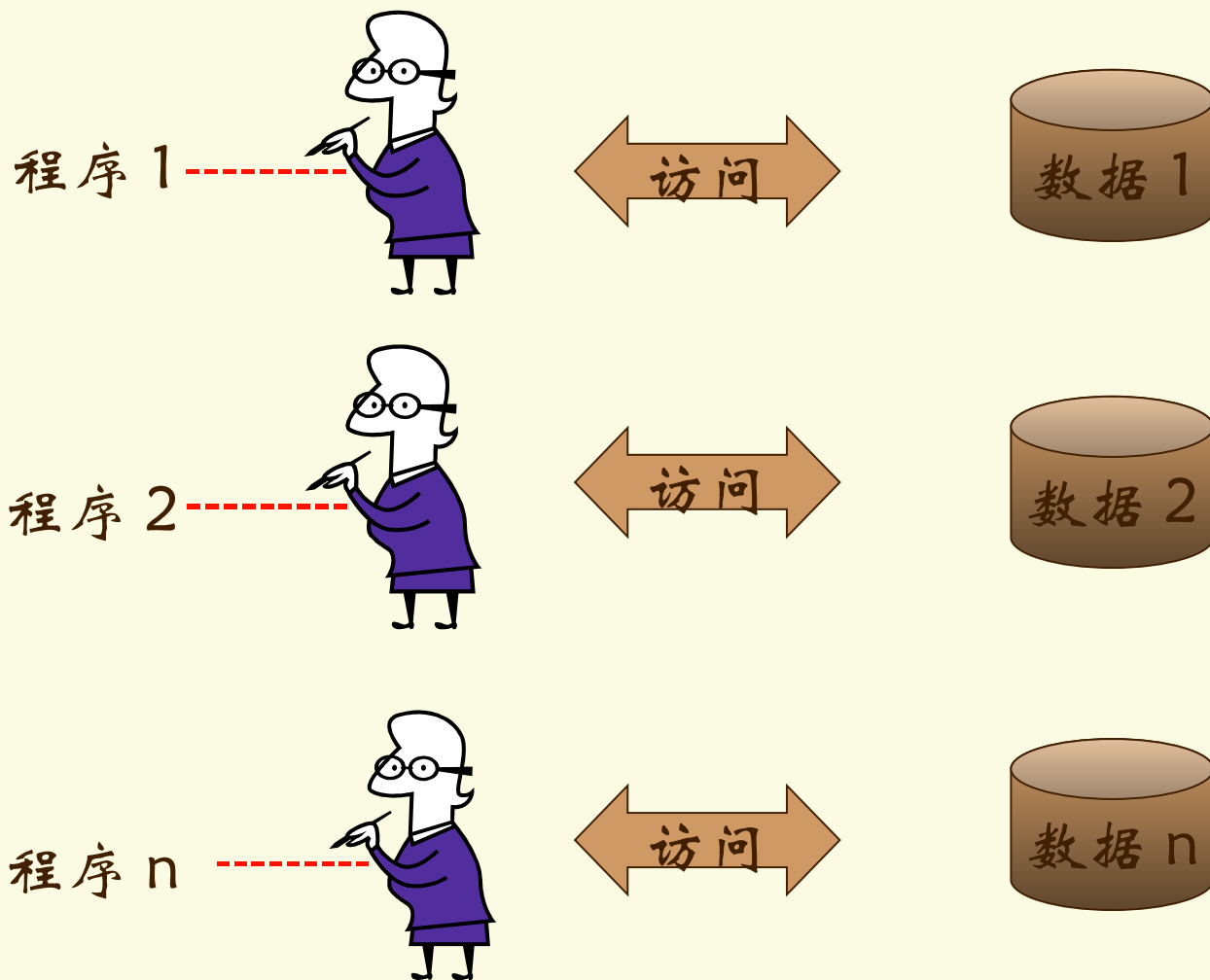


人工管理（续）

特点

- 数据的管理者：应用程序，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：数据和应用程序不独立
- 数据的结构化：无结构
- 数据控制能力应用程序自己控制

应用程序与数据的对应关系（人工管理）



二、文件系统

∞ 时期




- 50 年代末 --60 年代中

∞ 产生的背景

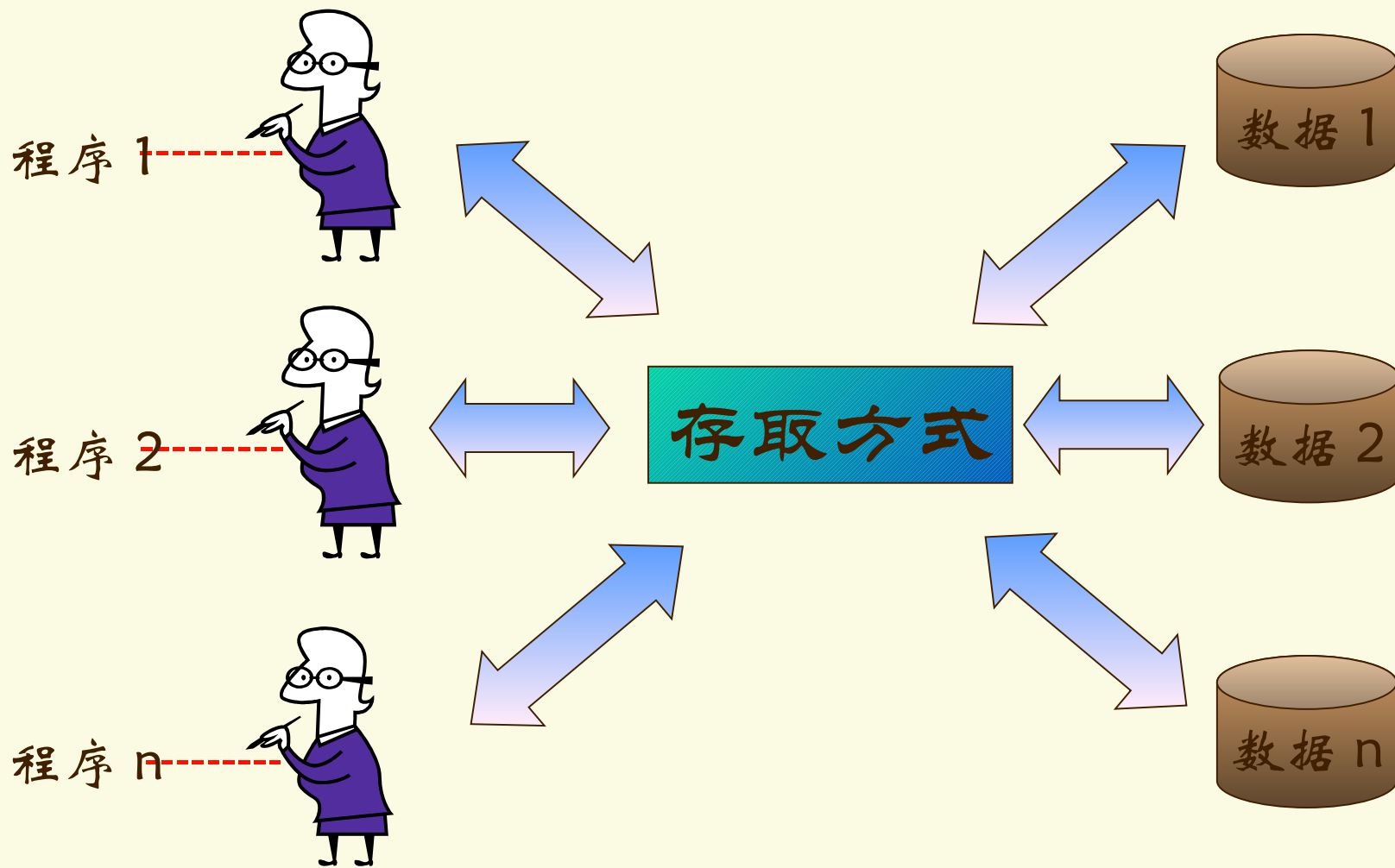
- | | |
|--------|------------|
| — 应用需求 | 科学计算、管理 |
| — 硬件水平 | 磁盘、磁鼓 |
| — 软件水平 | 有文件系统 |
| — 处理方式 | 联机实时处理、批处理 |

文件系统（续）

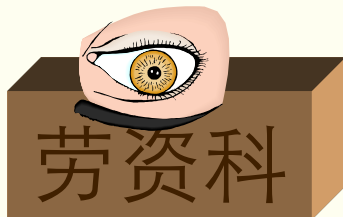
特点

- 数据的管理者：文件系统，数据可长期保存
 - “按文件名访问，记录存取”
- 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

应用程序与数据的对应关系（文件系统）

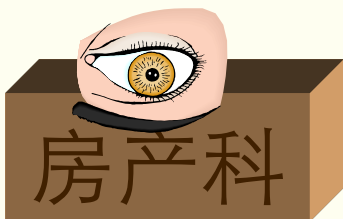


The Human Eye



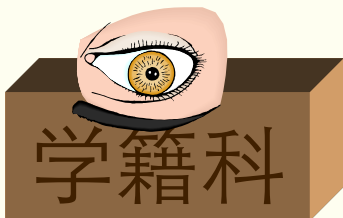
学号	姓名	系别	补贴
----	----	----	----

The Human Eye



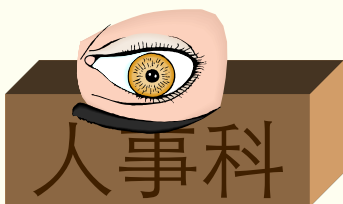
学号	姓名	性别	系别	住址
----	----	----	----	----

The Human Eye



学号	姓名	系别	学分	学位
----	----	----	----	----

The Human Eye



学号	姓名	性别	系别	年龄	学位	出身
----	----	----	----	----	----	----



文件系统中数据的数据结构

记录内有结构。

学生人事记录

学号	姓名	性别	系别	年龄	政治面貌	家庭出身	籍贯	家庭成员	奖惩情况
----	----	----	----	----	------	------	----	------	------

数据的结构是靠程序定义和解释的。

数据一般是定长的。

- 可以间接实现数据变长要求，但访问相应数据的应用程序复杂了。

学生人事主记录

学号	姓名	性别	系别	年龄	政治面貌	家庭出身	籍贯	家庭成员	奖惩情况
----	----	----	----	----	------	------	----	------	------

家庭成员记录

姓名	与本人关系	详细情况
----	-------	------

奖惩情况记录

日期	奖惩条目
----	------

900201	李明	男	计算机	24	共青团员	工人	江苏无锡		
--------	----	---	-----	----	------	----	------	--	--

家庭成员记录

李义山	父子
张 玉	母子
李 平	兄弟

奖惩情况记录

1991	一等奖学金
1991	吴玉章奖学金

文件系统中数据的结构

数据的最小存取单位是记录。

学生人事记录

学号	姓名	性别	系别	年龄	政治面貌	家庭出身	籍贯	家庭成员	奖惩情况
----	----	----	----	----	------	------	----	------	------



阿波罗登月计划

∞ 阿波罗飞船登月计划的需求

- 协调分散在全球制造的 200 万个阿波罗飞船零部件的生产进度
- 用文件系统开发了一个零部件生产计算机管理系统。
- 系统虽然可以工作，但由于文件系统分散管理的弱点，效率极低，60% 是冗余数据，维护十分困难。
- 该系统曾一度成为实现阿波罗计划的重大障碍之一

∞ 各国计算机学术界和工业界纷纷开展研究

- 数据建模、数据模型研究与实现的探索
- 成果是出现了一种全新的高效的数据管理技术——数据库技术

三、数据库系统

∞ 时期

- 60 年代末以来

∞ 产生的背景

- | | |
|--------|-----------------|
| — 应用背景 | 大规模管理 |
| — 硬件背景 | 大容量磁盘 |
| — 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 |

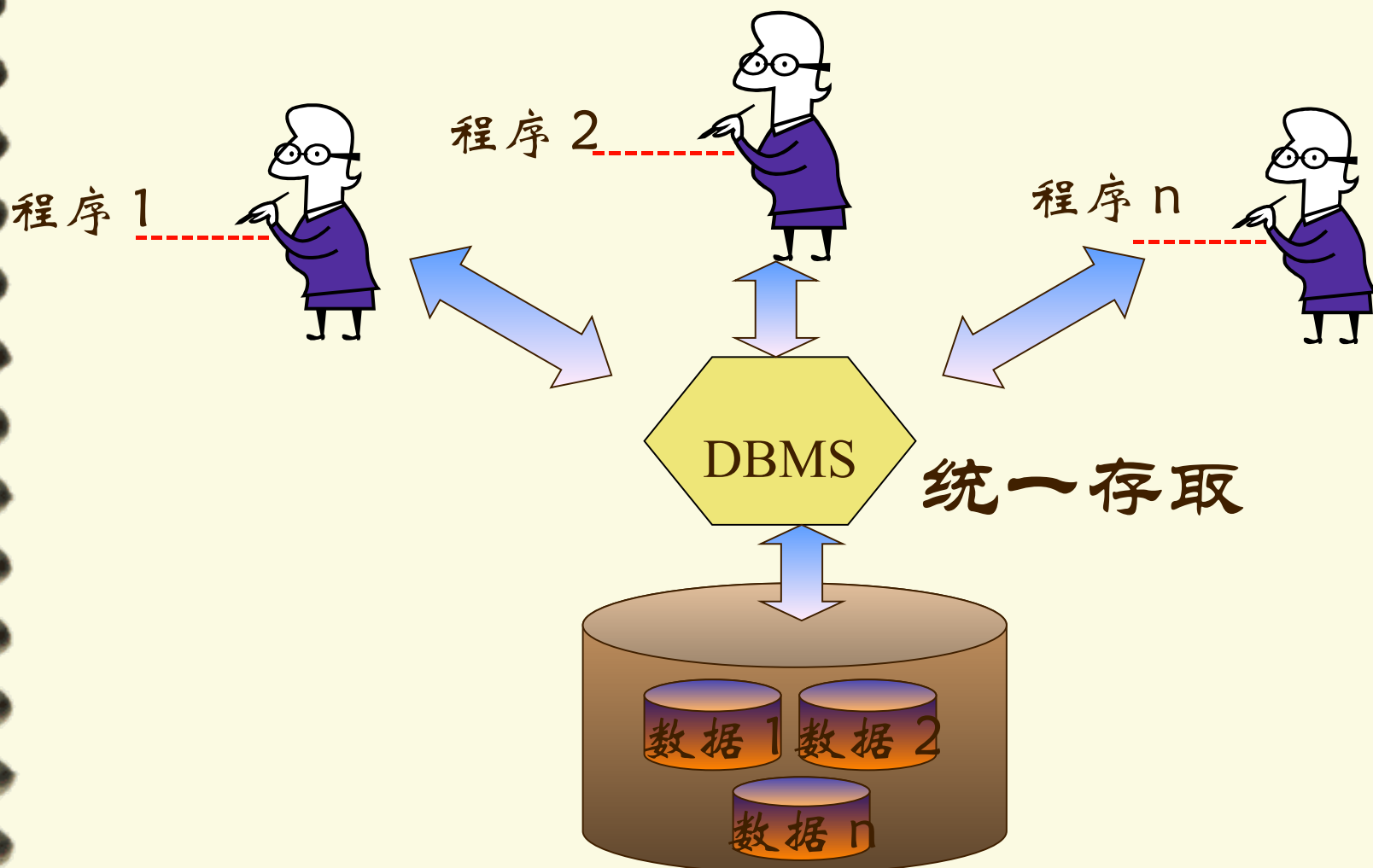
数据库系统（续）

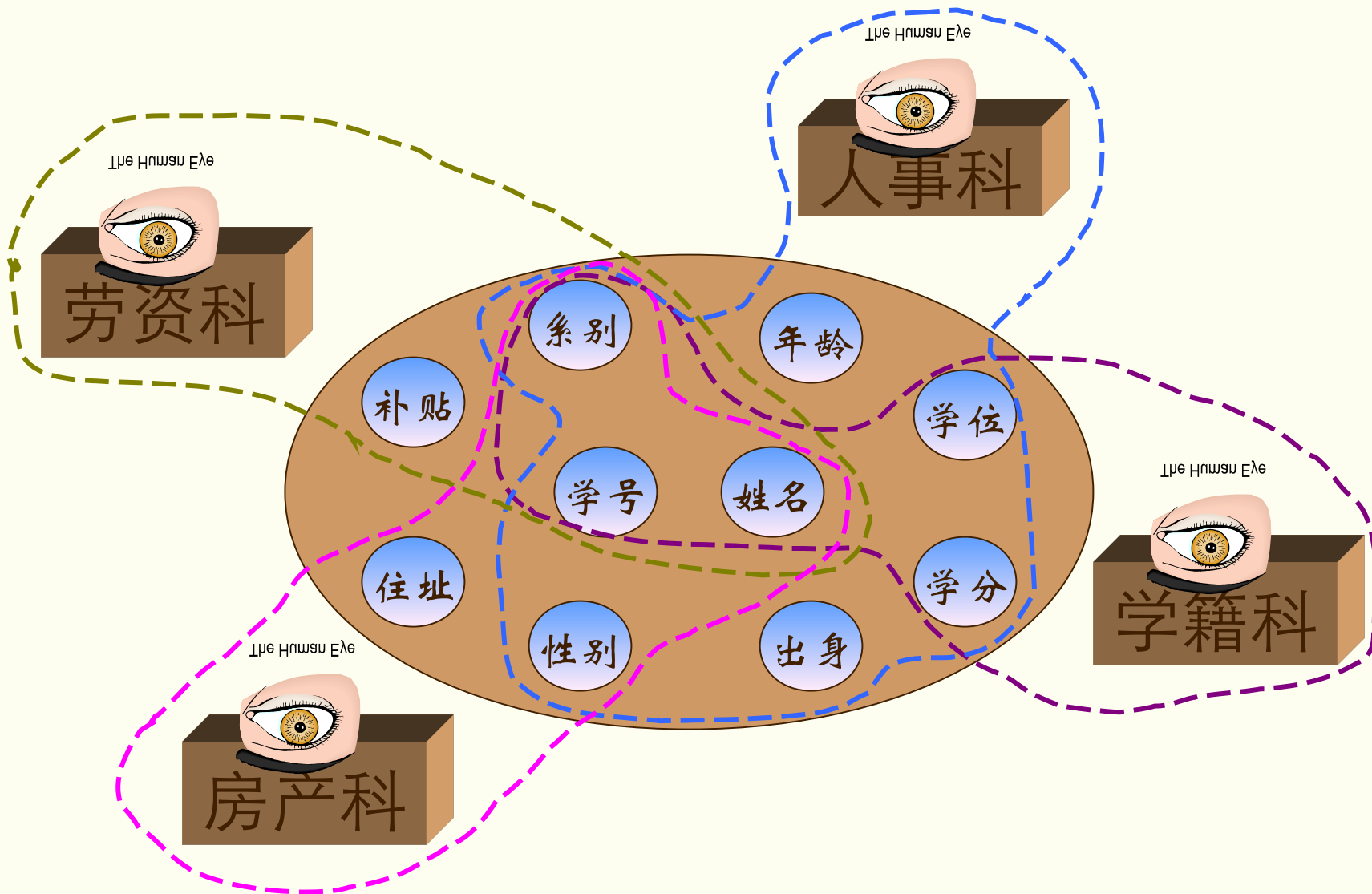
特点

- 数据的管理者：DBMS
- 数据面向的对象：现实世界 
- 数据的共享程度：共享性高 
- 数据的独立性：高度的物理独立性和一定的
逻辑独立性 

- 数据的结构化：整体结构化 
- 数据控制能力：由 DBMS 统一管理和控制

应用程序与数据的对应关系（数据库系统）





数据的高共享性的好处

- ∞ 降低数据的冗余度，节省存储空间
- ∞ 避免数据间的不一致性
- ∞ 使系统易于扩充



数据独立性

∞ 数据独立性

- 数据的物理独立性
- 数据的逻辑独立性

∞ 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。

数据独立性（续）

∞ 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。



数据结构化

- ∞ 整体数据的结构化是数据库的主要特征之一。
- ∞ 数据库中实现的是数据的真正结构化
 - 数据的结构用数据模型描述，无需程序定义和解释。
 - 数据可以变长。
 - 数据的最小存取单位是数据项。



学生基本记录

学号	姓名	性别	系别	年龄
----	----	----	----	----

学生人事记录

家庭出身	籍贯	政治面貌
------	----	------

奖惩情况记录

日期	奖惩条目
----	------

家庭成员记录

姓名	与本人关系	详细情况
----	-------	------

学生学籍记录

日期	地点	学历名
----	----	-----

学生选课记录

课程号	学号	学分
-----	----	----

课程记录

课程号	课程名
-----	-----



DBMS 对数据的控制功能

☞ 数据的安全性（Security）保护

- 使每个用户只能按指定方式使用和处理指定数据，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。

DBMS 对数据的控制功能

∞ 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。





長白山天池