

管理信息系统

2009.2基于DB的信息系统
D 数据库和E-R图

数据库和E-R图

- 关系数据库：数据的逻辑结构
- E-R图：逻辑的表达

回顾：用EXCEL管理数据的弱点

- 1 数据冗余，并且容易带来数据错误
- 2 难于交流共享，难于统一管理
- 3 基于简单的文件管理方式
 - 不能多人同时读写，难于网络共享
 - 多复制几份，得到的人各做各的修改，版本混乱
- 解决：数据库

管理名称	学号	姓名	性别	年龄	成绩	课程名称	课程	管理名称
管理信息系统	200511101001	张三	男	20	80	2008 (2)	(二) 214	色彩理论
管理信息系统	200511101002	李小明	男	21	48	75	2008 (2)	(二) 214
管理信息系统	200511101003	王小二	男	20	48	88	2008 (2)	(二) 214
管理信息系统	200511201001	赵小童	女	20	48	90	2008 (2)	(二) 214
管理信息系统	200511201002	孙小红	女	19	48	69	2008 (2)	(二) 214
管理信息系统	200511201003	张三	男	22	48	72	2008 (2)	(二) 214
数据库	200511101001	张三	男	20	32	81	2007 (1)	(一) 101
数据库	200511101002	李小明	男	21	32	92	2007 (1)	(一) 101
数据库	200511101003	王小二	男	20	32	80	2007 (1)	(一) 101
数据库	200511201001	赵小童	女	20	32	77	2007 (1)	(一) 101
数据库	200511201003	张三	男	22	32	65	2007 (1)	(一) 101

关系数据库

• 拆分、Access 体会

“关系数据库”

“数据的逻辑联系”

什么是数据库

- 相关概念、教材P103 (第6版)

数据库 包含：

- 数据的具体内容** ← 光有数据并不够
- 数据的逻辑结构** ← 只有严格遵守这些约束的数据，才被允许写入

关系数据库

- 数据库模型：关系型的数据库是主流应用
 - (知道这几个名词就可以)
 - 文件型
 - 层次型
 - 关系型
 - 面向对象
- 未来
 - 数据集市，数据仓库[技术]
 - XML [技术]
 - 数据集成，数据资源管理[管理]

报表例子：概览

表：类别

类别ID	类别名称	说明
1	饮料	软饮料、咖啡、茶、啤酒和酒精类
2	调味品	食用盐和胡椒粉、香料、烹饪调味品
3	点心	饼干、糖果和甜食
4	日用品	洗洁精
5	宠物食品	猫粮、狗粮、宠物训练用品
6	肉/家禽	猪肉的
7	海鲜类	干海鲜和鱼类
8	海鲜	海产品类

表：产品

产品ID	产品名称	规格	单位	单位数量	单价	原零售	停止
1	牛奶	任意	1	每箱24瓶	¥18.00	35	Yes
2	牛奶	任意	1	每箱24瓶	¥18.00	17	No
3	牛奶	任意	1	每箱24瓶	¥18.00	13	No
4	肉	任意食品	1	每箱12瓶	¥22.00	25	No
5	肉	任意食品	2	每箱12瓶	¥21.25	0	Yes
6	肉	任意	2	每箱12瓶	¥21.00	120	No
7	肉	任意	7	每箱24瓶	¥38.00	12	No

最终报表：各国雇员销售额

查询：各国雇员销售额

国家	雇员	名字	实际日期	订单ID	销售金额
中国	周	周强	1997-01-01	10000	¥1,315.00
中国	王	王华	1997-01-01	10001	¥1,446.00
中国	张	张强	1997-01-01	10002	¥2,486.00
中国	陈	陈强	1997-01-01	10003	¥1,442.00
中国	沙	沙强	1997-01-01	10004	¥2,122.00

查询：订单小计

订单ID	小计
10000	¥1,315.00
10001	¥1,446.00
10002	¥2,486.00
10003	¥1,442.00
10004	¥2,122.00

表：雇员

雇员ID	姓名	职务	雇用日期
1	周	销售经理	1997-01-01
2	王	销售经理	1997-01-01
3	张	销售经理	1997-01-01
4	陈	销售经理	1997-01-01

表：订单

订单ID	客户	雇员ID	日期	地区	运费
10000	山东金泰	7	1996-07-04	德州	¥32.30
10001	山东金泰	7	1996-07-04	德州	¥32.30
10002	山东金泰	7	1996-07-04	德州	¥32.30
10003	山东金泰	7	1996-07-04	德州	¥32.30
10004	山东金泰	7	1996-07-04	德州	¥32.30

数据的逻辑关系

- 一张二维表，描述一个对象（实体）
 - 对象（实体）：客观世界中的人、部门或事物等
 - 每一列记载了该实体的一个属性，称“字段”
 - 每一行记载了具体的一个对象，称“记录”
- 每一个字段，都有自己的属性、长度约束
- “主关键字”：“KEY”
 - 唯一的编号，可以区别每一行记录
 - 不能为空值

数据的逻辑关系

- 表之间的“逻辑关系”
- 1 两个表都有自己的主关键字
- 2 表A的“某字段” + 表B的“主关键字”形成一对关系
 - 数据可以重复
 - 数据不允许重复
 - “外部关键字”

数据的逻辑关系

- 表“关系”
 - 1: N
 - 1: 1
- “关系”形成后会有“完整性约束”

更多的例子

表：客户

客户ID	公司名称	联系人姓名	联系人电话	地址	城市	邮政编码	电话
ALFKI	三河实业有限公司	周小姐	销售经理	大亚湾路 50 号	天津	345007	(020) 30074211
ANATK	天津泰来公司	王先生	销售	承德路 80 号	天津	234015	(020) 3154729
ANTON	天津有限公司	李先生	销售经理	天津路 20 号	天津	090079	(020) 40857789
BLAND	天津有限公司	李先生	销售经理	天津路 20 号	天津	790009	(020) 91234567

表：订单

订单ID	客户ID	雇员ID	日期	地区	运费
10000	ALFKI	周小姐	1996-07-04	德州	¥32.30
10001	ANATK	王先生	1996-07-05	德州	¥11.05
10002	ANTON	李先生	1996-07-06	德州	¥65.83
10003	BLAND	李先生	1996-07-08	德州	¥41.34

表：订单明细

订单ID	产品	单价	数量	折扣
10000	饼干	¥14.00	12	0.00%
10000	饼干	¥9.00	10	0.00%
10000	饼干	¥16.00	5	0.00%
10000	饼干	¥18.00	8	0.00%

表：产品

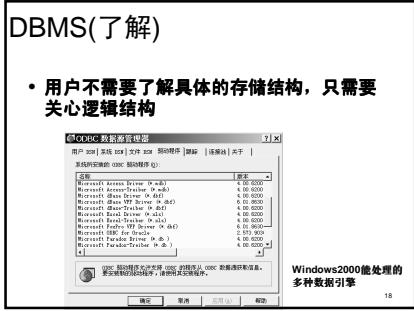
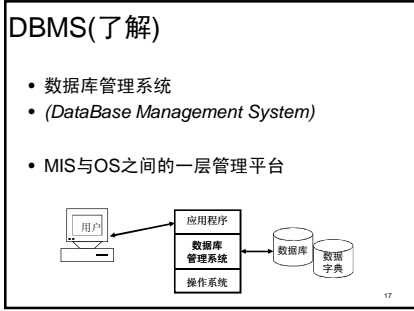
产品ID	产品名称	供应商	类别	单位数量	单价	库存量	中止
1	饼干	供应商	饼干	每包24包	¥14.00	20	Yes
2	饼干	供应商	饼干	每包24包	¥9.00	17	No
3	饼干	供应商	饼干	每包24包	¥16.00	13	No
4	饼干	供应商	饼干	每包24包	¥18.00	53	No

数据的逻辑关系

- 数据“关系”的作用
- 1 确保数据正确
 - 完整性约束
- 2 在查询时，确保正确连接各个表中的数据，得到正确的结果
- 在上机实验中进一步体会

总结：关系数据库的特点

- 原始数据保存在若干二维“表”中，每个表包括：
 - 字段定义，主关键字，若干行数据
- 表之间有严格的关联定义
 - 减少冗余、方便查询
- 用户使用数据时，较少面对表，较多面对查询
 - 查询将各个表的数据联系起来呈现给用户
 - 录入界面、查询界面、报表的基础通常都是查询



DBMS(了解)

- SQL 语句

SELECT: 实现数据库查询	UPDATE: 更新原来的记录
CREATE: 定义新的数据库	DELETE: 删除原来的记录
INSERT: 插入新记录	GRANT: 实现数据控制

19

关系数据库小结

- 实现的焦点
 - 表的设计 (E-R图分析, 接下来的内容)
 - 查询的设计 (实验课程内容)
 - 界面的设计 (实验课程内容)
- 设计的焦点
 - 怎么能保证得到的系统是管理者想要的呢? (后续课程的内容)

20

数据库和E-R图

- 关系数据库: 数据的逻辑结构
- E-R图: 逻辑的表达

21

数据关系模型

- E-R图:
 - 一种逻辑分析工具, 用来得到正确的数据关系

22

E-R图

- E-R图 (实体-联系图):
 - 描述现实世界的概念模型
- 实体: **实体**
- 属性: **属性**
- 联系:
 - 一对一联系
 - 一对多联系
 - 多对多联系 -- 本课程不允许出现

24

E-R图

- 实体 (Entity):
 - 是一个有着一系列显著的、易辨认的属性的对象。
 - 实体可以是具体的 (物体、人物等)
 - 实体也可以抽象的 (事、概念、事物之间的联系)
 - 学生, 教师, 上课, 考核

25

E-R图

- 属性 (Attribute)
 - 实体的特性, 它描述了实体的一个部分。
 - 一个实体可由若干个属性来刻画。
 - 如学生 (学号, 姓名, 性别,)
- 关键字 (Key):
 - 一个属性, 能区别每个独一无二的实体。
 - 如: 学号为学生的关键字。

26

E-R图

学号	课程号	成绩	任课教师
20051101001	1	85	
20051101001	2	81	
20051101002	1	75	
20051101002	2	92	
20051101003	1	88	
20051101003	2	80	
200511201001	1	90	
200511201001	2	77	
200511201002	1	83	
200511201003	1	72	
200511201003	2	65	

27

E-R图

28

E-R图

- 分析“学生”
- 学号、姓名、性别、年龄
- 加入“入学日期”和“专业”是否合适？

冗余

E-R图

- 加入新的对象“班级”

E-R图

- 分析“课程”
- 课程编号、名称、教室、学时数
- 加入“课程介绍”、“教学大纲”是否合适？

E-R图

- 加入新对象“课程计划”

E-R图：完善

E-R图：模型的扩展

E-R图的规范化（了解）

- 分析数据的一些逻辑判断标准
- 观察数据的几个角度
 - 数据冗余
 - 数据异常：修改异常、插入异常、删除异常
- 5个范式

规范化（了解）

零件号	仓库号	零件数量	仓库地址
1	1	100	鼓楼街
2	1	150	鼓楼街
3	1	100	鼓楼街
4	2	150	火车站

观察冗余：“仓库地址”重复了多次

观察异常：
 • 当某个零件换仓库时，必须改变仓库号，同时不能忘记修改仓库地址
 • 当新增某个零件时，也必须保证加入的仓库号和仓库地址匹配无误

观察到了冗余和异常，说明这个表需要拆分

规范化（了解）

- 第一规范型：字段切割到最细
 - 如果一个关系模式中的属性都是单纯的（即不可再分为更小的属性），则称该模式是属于第一规范型（First Normal Form, 即 1NF）。

姓名	地址	邮编
甲	江苏省南京市卫岗	210095

↓

姓名	省	市	街道	邮编
甲	江苏	南京	卫岗	210095

规范化（了解）

- 第二规范型：关键字与其他字段的关系
 - 如果一个关系模式属于1NF，并且所有的非关键字都完全地依赖于关键字（即不存在部分依赖），则称该关系模式属于第二规范型，即2NF。

只依赖于一个关键字

(KEY)	学号	姓名	课程	成绩
99001	Lily	Lily	C语言	91
99001	Lily	Lily	数据结构	82
99002	Keven	Keven	C语言	77
99002	Keven	Keven	数据结构	86

同时依赖于两个关键字

(KEY)	学号	姓名	课程	成绩
99001	Lily	Lily	C语言	91
99001	Lily	Lily	数据结构	82
99002	Keven	Keven	C语言	77
99002	Keven	Keven	数据结构	86

规范化（了解）

- 第二规范型

(KEY)	学号	姓名	课程	成绩
99001	Lily	Lily	C语言	91
99001	Lily	Lily	数据结构	82
99002	Keven	Keven	C语言	77
99002	Keven	Keven	数据结构	86

(KEY)	学号	姓名
99001	Lily	Lily
99002	Keven	Keven

规范化（了解）

- 第三规范型：非关键字之间的关系
 - 如果一个关系模式属于2NF，并且不存在非关键字传递地依赖于关键字，则称该关系模式属于第三规范型（3NF）。

(KEY)	零件号	仓库号	零件数量	仓库地址
1	1	1	100	鼓楼街
2	1	1	150	鼓楼街
3	1	1	100	鼓楼街
4	2	2	150	火车站

规范化（了解）

- 第三规范型

(KEY)	零件号	仓库号	零件数量	仓库地址
1	1	1	100	鼓楼街
2	1	1	150	鼓楼街
3	1	1	100	鼓楼街
4	2	2	150	火车站

(KEY)	零件号	仓库号	零件数量
1	1	1	100
2	1	1	150
3	1	1	100
4	2	2	150

(KEY)	仓库号	仓库地址
1	1	鼓楼街
2	2	火车站

规范化（了解）

- 第四规范型（不一定要遵守）
 - 如果一个关系模式属于3NF，并且至多只含一个多值依赖，则称该关系模式属于第四规范型（4NF）。
- 第五规范型（不一定要遵守）
 - 如果一个关系模式属于4NF，并且不能被三个以上的较小的关系所替代，则称该关系模式为第五规范型（5NF）。

规范化（了解）

- 基于4NF、5NF优化设计（了解）

(KEY)	翻译	方言	外语
王兵	上海话	英语	
王兵	广东话	英语	
王兵	福建话	英语	
王兵	上海话	法语	
王兵	广东话	法语	
王兵	福建话	法语	
王兵	上海话	德语	
王兵	广东话	德语	
王兵	福建话	德语	

正常：需要翻译英文资料，可以找王同志；需要翻译广东话资料，也可以找王同志

异常：当发现王同志的德语不过关，不能再交给她相关任务时，必须保证同时删除与“德语”对应的三条记录。

当发现王同志两班牙文特长时，要增加记录，也必须同时增加三条。

规范化（了解）

- 基于4NF、5NF优化设计

(KEY)	翻译	方言	外语
王兵	上海话	英语	
王兵	广东话	英语	
王兵	福建话	英语	
王兵	上海话	法语	
王兵	广东话	法语	
王兵	福建话	法语	
王兵	上海话	德语	
王兵	广东话	德语	
王兵	福建话	德语	

(KEY)	翻译	外语
王兵	英语	
王兵	法语	
王兵	德语	

(KEY)	翻译	方言
王兵	上海话	
王兵	广东话	
王兵	福建话	

规范化：守则

- 适当的规范化
 - 数据冗余 ↓ 更强壮
 - 更新异常 ↓ 效率降低
 - 连接运算时间 ↑
 - 查询时间 ↑

规范化程度高

(KEY)	翻译	分类	能处理的语种
王兵	外语	英语	
王兵	外语	法语	
王兵	外语	德语	
王兵	中国方言	上海话	
王兵	中国方言	广东话	
王兵	中国方言	福建话	

比起前一页的表设计，这个设计不那么规范，有冗余，可以接受

添加或者删除一门语言，只需要操作一行数据

E-R图和数据规范化：小结

- 确保数据模型正确转换的技术
- E-R图：发掘对象的属性集和关系集
 - 决定需要哪些表
 - 决定每个表需要哪些字段
 - 决定表之间的关系
- 规范化：确定表的正确分割
 - 指出表中存在的冗余，给出解决方法