课程名称: 数据库系统

第11讲:(第8章)

关系模式设计优化

单 位：重庆大学计算机学院

1.关系模型的好与坏

- 假设有如下关系模式 (简称 SNLRHW)
  Hourly_Emps(*ssn, name, lot, rating, hrly_wage, hrs_worked)*
- 如果其中rating 决定 hrly_wage，这个关系模式有没有问题?

S	N	L	R	W	H
123-22-3666	Attishoo	48	8	10	40
231-31-5368	Smiley	22	8	10	30
131-24-3650	Smethurst	35	5	7	30
434-26-3751	Guldu	35	5	7	32
612-67-4134	Madayan	35	8	10	40

潜在的问题

- *能否只修改*SNLRWH的第一元组的w属性的值？

——修改异常

- *如果想添加一个员工信息又不知道他的等级所对应的时薪，会如何？*

——添加异常

- 如果要删除所有等级为5的员工信息，会如何？
- 会丢失5级员工时薪信息
- *——删除异常*

冗余

- 当数据的某些部分能由其他部分推导出来，就意味着存在冗余。
  – 例如: Smiley的 hrly_wage 可以从Attishooe 的 hrly_wage推出，因为他们等级相同且等级rating决定hrly_wage。
- 冗余的存在是因为存在完整性约束 *(e.g., FD: R→ W).*

如何解决冗余问题

- 由于存在约束，特别是函数依赖，导致冗余的产生。
- 如果存在冗余，就需要对关系模式进行优化。

– 主要优化技术: 模式分解 *(比如,* ABCD 用 AB和BCD, 或者 ACD 和 ABD替代)。

- 如何正确地使用模式分解:

– 是否有必要分解一个模式?

– 分解以后会不会出现什么问题?

分解

S	N	L	R	W	H
123-22-3666	Attishoo	48	8	10	40
231-31-5368	Smiley	22	8	10	30
131-24-3650	Smethurst	35	5	7	30
434-26-3751	Guldu	35	5	7	32
612-67-4134	Madayan	35	8	10	40

=

S	N	L	R	H
123-22-3666	Attishoo	48	8	40
231-31-5368	Smiley	22	8	30
131-24-3650	Smethurst	35	5	30
434-26-3751	Guldu	35	5	32
612-67-4134	Madayan	35	8	40

$\bowtie$

R	W
8	10
5	7

良好关系设计的特点

关系更新异常的实例(联系集开课R)

学生学号 S#	课程编号 C#	课程名称 CN	学生成绩 G	学生姓名 TN	开课系名 D	← 注释
001	J01	C++	90	张三	计算机系	
002	J01	C++	96	张三	计算机系	
003	J01	C++	100	张三	计算机系	
004	J01	C++	78	张三	计算机系	
001	J02	Java	65	李四	计算机系	
002	J02	Java	59	李四	计算机系	
003	J02	Java	85	李四	计算机系	
001	S01	数学分析	85	王五	数学系	
002	W01	大学物理	60	菲菲	物理系	

修改异常：当改变课程J01的信息（C#,CN,D）时，4个相应记录的值都得一同改变。

插入异常：不能插入未开课程的课程记录。因为：无学生信息，而主码(S#,C#）不能为空。

删除异常：当所有学生退选J02课，则课程J02的信息也会消失(一同被删除)。

良好关系设计的特点

- *合并instructor和department 为inst_dept会怎么样?*

instructor

ID	*name*	*dept_name*	*salary*
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

department

dept_name	*building*	*budget*
Biology	Watson	90000
Comp. Sci.	Taylor	100000
Elec. Eng.	Taylor	85000
Finance	Painter	120000
History	Painter	50000
Music	Packard	80000
Physics	Watson	70000

Instr_dept

ID	*name*	*salary*	*dept_name*	*building*	*budget*
22222	Einstein	95000	Physics	Watson	70000
12121	Wu	90000	Finance	Painter	120000
32343	El Said	60000	History	Painter	50000
45565	Katz	75000	Comp. Sci.	Taylor	100000
98345	Kim	80000	Elec. Eng.	Taylor	85000
76766	Crick	72000	Biology	Watson	90000
10101	Srinivasan	65000	Comp. Sci.	Taylor	100000
58583	Califieri	62000	History	Painter	50000
83821	Brandt	92000	Comp. Sci.	Taylor	100000
15151	Mozart	40000	Music	Packard	80000
33456	Gold	87000	Physics	Watson	70000
76543	Singh	80000	Finance	Painter	120000

部门表信息会大量重复!

2.函数依赖(FDs)

- 函数依赖可形式化表示为: X→Y, 其中 X and Y 是属性集。
- 例如: rating→hrly_wage，AB →C
- 如果对于关系实例r中的任意一对元组t1和t2，有 *t1.X = t2.X 逻辑蕴涵 t1.Y =t2.Y，那么，函数依赖* X→Y在关系r中成立。
- 即,关系r中给定两个元组，如果 *X 属性值相等, 则 Y 的值也必须相等.* (X 和 Y 是属性集）

函数依赖(FDs)

- *如果对关系R中的每个实例r，都满足函数依赖，则该函数依赖在关系R上成立。*
  - 函数依赖必须由应用的语义所决定
  - 对于给定关系R的某实例r，我们能判断它是否违反了某个函数依赖。
  - 但是不能仅仅根据一个满足函数依赖的实例来断定该函数依赖在关系模式R上成立。

例: 实体集上的约束

- 考虑以下关系Hourly_Emps:
- Hourly_Emps (*ssn, name, lot, rating, hrly_wage, hrs_worked)*
- 在该关系上成立的函数依赖:
- *ssn 是主码: S → SNLRWH*
- *rating 决定 hrly_wage: R → W*

A	B	C
1	1	2
1	1	3
2	1	3
2	1	2

在该关系实例上存在哪些函数依赖？

FDs with A as the left side:	**Satisfied by the relation instance?**
$A \rightarrow A$	yes
$A \rightarrow B$	yes
$A \rightarrow C$	No
$A \rightarrow AB$	yes
$A \rightarrow AC$	No
$A \rightarrow BC$	No
$A \rightarrow ABC$	No

A	B	C
1	1	2
1	1	3
2	1	3
2	1	2

FD	**Satisfied by the relation instance?**
$C \rightarrow B$	yes
$C \rightarrow AB$	No
$B \rightarrow C$	No
$B \rightarrow B$	Yes
$AC \rightarrow B$	Yes [note!]
…	…

函数依赖理论

- 如何从一个给定的函数依赖集推导出它所逻辑蕴涵的所有函数依赖呢？
- Armstrong 公理 (X, Y, Z 是属性集):

– *Reflexivity*（自反律）:

If $X \supseteq Y$, then $X \rightarrow Y$

– *Augmentation*（增广律）:

If $X \rightarrow Y$, then $XZ \rightarrow YZ$ for any Z

– *Transitivity*（传递律）:

If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Armstrong 公理

- 由Armstrong公理导出的推理规则
  - 合并律(union rule)
    - 若$X \to Y$，$X \to Z$，则$X \to YZ$
  - 分解律(decomposition rule)
    - 若$X \to YZ$，则$X \to Y$，$X \to Z$
  - 伪传递律(pseudotransitivity rule)
    - 若$X \to Y$，$WY \to Z$，则$WX \to Z$

例

- $R = (A, B, C, G, H, I)$
  $F = \{A \rightarrow B; A \rightarrow C; CG \rightarrow H; CG \rightarrow I; B \rightarrow H\}$
- 请推导出以下函数依赖?
  - $A \rightarrow H$
  - $AG \rightarrow I$
  - $CG \rightarrow HI$