

数据库系统之三

--数据建模与数据库设计



第14讲 函数依赖及其公理/定理

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on Intelligent
Computing for Enterprises & Services,
Harbin Institute of Technology

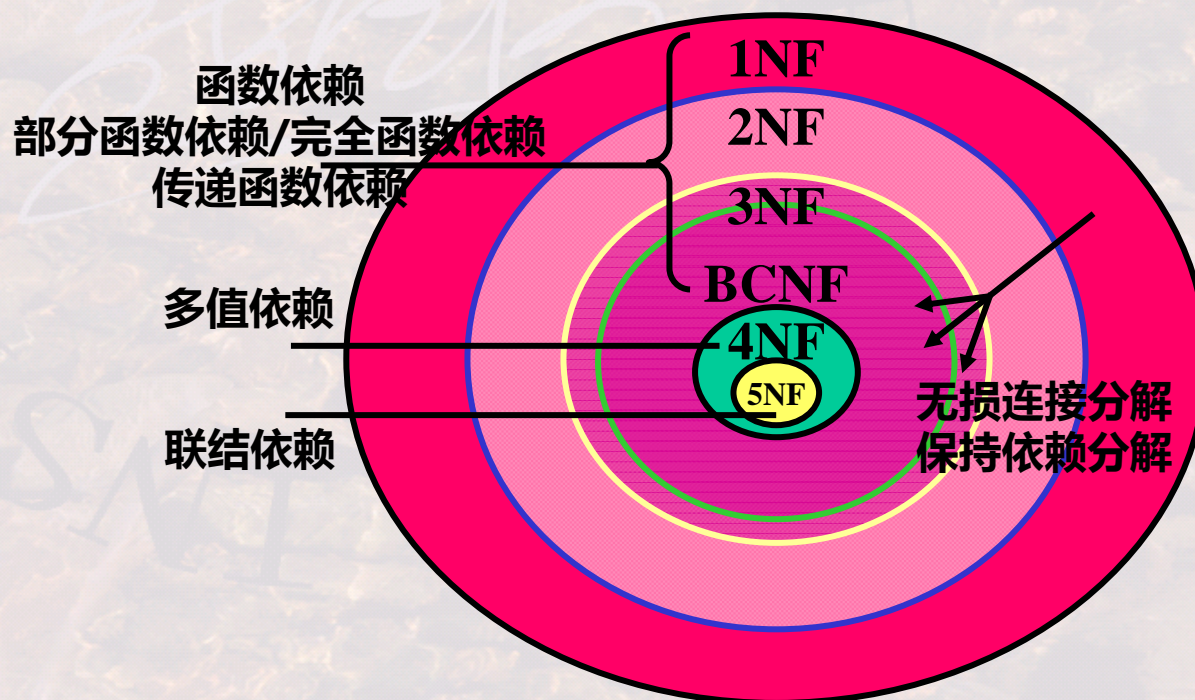
本讲学习什么？

如何避免数据库的一致性问题——数据库的规范性设计

数据库的规范性设计需要分析数据库Table中的属性在取值方面有什么依存关系？数据库设计过程中应遵循什么样的原则

➤ 数据库设计理论

- 数据依赖理论
- 关系范式理论
- 模式分解理论



本讲学习什么？



基本内容

1. 函数依赖
2. 完全函数依赖与传递函数依赖
3. 关于函数依赖的公理和定理
4. 函数依赖集的最小覆盖

重点与难点

- 一组概念：函数依赖、部分函数依赖和完全函数依赖、传递函数依赖、候选键、非主属性、逻辑蕴涵、闭包、属性闭包、覆盖、最小覆盖等
- 关于函数依赖的公理和定理，相关的证明
- 求属性闭包的算法、求最小覆盖的算法

函数依赖

战德臣

哈尔滨工业大学 教授.博士生导师
黑龙江省教学名师
教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

[Definition] 函数依赖

设 $R(U)$ 是属性集合 $U=\{A_1, A_2, \dots, A_n\}$ 上的一个关系模式， X, Y 是 U 上的两个子集，若对 $R(U)$ 的任意一个可能的关系 r ， r 中不可能有两个元组满足在 X 中的属性值相等而在 Y 中的属性值不等，则称“ X 函数决定 Y ”或“ Y 函数依赖于 X ”，记作 $X \rightarrow Y$ 。

示例： $U = \{\text{学号}, \text{姓名}, \text{年龄}, \text{班号}, \text{班长}, \text{课号}, \text{成绩}\}$

- 学号 \rightarrow { 姓名, 年龄}
- 班号 \rightarrow 班长
- { 学号, 课号} \rightarrow 成绩

设计关系模式时，除给出属性全集外，还需给出数据依赖集合

注：函数依赖的分析取决于对问题领域的限定和分析，取决于对业务规则的正确理解。例如：问题领域中，学生是没有重名的，则有：“年龄”和“家庭住址”都函数依赖于“姓名”。而在另一个问题领域中，学生是有重名的，则上述函数依赖是不成立的。

函数依赖

(2)函数依赖的示例

示例：下表就是问题领域, 则存在的函数依赖有哪些呢?

属性A	属性B	属性C
1	2	3
4	2	3
5	3	3

属性A	属性B	属性C	属性D
a1	b1	c1	d1
a1	b2	c1	d2
a2	b2	c2	d2
a2	b3	c2	d3
a3	b3	c2	d4

函数依赖

(2)函数依赖的示例

示例：下表就是问题领域, 则存在的函数依赖有哪些呢?

✓ 下表存在的函数依赖有: $A \rightarrow B$, $B \rightarrow C$

属性A	属性B	属性C
1	2	3
4	2	3
5	3	3

✓ 下表存在的函数依赖有: $A \rightarrow C$, $D \rightarrow B$

属性A	属性B	属性C	属性D
a1	b1	c1	d1
a1	b2	c1	d2
a2	b2	c2	d2
a2	b3	c2	d3
a3	b3	c2	d4

函数依赖的特性

- (1)对 $X \rightarrow Y$, 但 $Y \not\subseteq X$, 则称 $X \rightarrow Y$ 为**非平凡的函数依赖** ;
- (2)若 $X \rightarrow Y$, 则任意两个元组 , 若 X 上值相等 , 则 Y 上值必然相等,则称 X 为决定因素 ;
- (3)若 $X \rightarrow Y$, $Y \rightarrow X$, 则记作 $X \leftrightarrow Y$;
- (4)若 Y 不函数依赖于 X , 则记作 $X \not\rightarrow Y$;
- (5) $X \rightarrow Y$, 有基于模式 R 的 , 则要求对任意的关系 r 成立 ; 有基于具体关系 r 的 , 则要求对某一关系 r 成立 ;
- (6)如一关系 r 的某属性集 X , r 中根本没有 X 上相等的两个元组存在 , 则 $X \rightarrow Y$ 恒成立 ;

练习：请分析下列属性集上的函数依赖

- 学生(学号, 姓名, 班级, 课号, 课程名, 成绩, 教师, 教师职务)
- 员工(员工码, 姓名, 出生日期, 联系电话, 最后学历, 毕业学校, 培训日期, 培训内容, 职务变动日期, 变动后职务)
- 图书(书号, 书名, 出版日期, 出版社, 书架号, 房间号)
- 客户(客户号, 客户名称, 类别, 联系电话, 产品编码, 产品名称, 数量, 要货日期)

练习：请分析下列属性集上的函数依赖

- 学生(学号, 姓名, 班级, 课号, 课程名, 成绩, 教师, 教师职务)

✓学号 \rightarrow {姓名, 班级}; 课号 \rightarrow 课程名; {学号, 课号} \rightarrow 成绩

✓教师 \rightarrow 教师职务

✓{班级, 课号} \rightarrow 教师

{班级, 课号} \rightarrow 教师;

课号 \rightarrow 教师;

{学号, 课号} \rightarrow 教师

究竟选哪一个取决于对问题领域的理解

- 客户(客户号, 客户名称, 类别, 联系电话, 产品编码, 产品名称, 数量, 要货日期)

✓客户号 \rightarrow {客户名称, 类别}

✓产品编码 \rightarrow 产品名称

✓{客户号, 产品编码, 要货日期} \rightarrow 数量

本质上,函数依赖
是对属性之间取值
的一种约束,
是一种数据依赖

完全函数依赖与传递函数依赖

战德臣

哈尔滨工业大学 教授.博士生导师
黑龙江省教学名师
教育部大学计算机课程教学指导委员会委员

Research Center on Intelligent
Computing for Enterprises & Services,
Harbin Institute of Technology

[Definition] 部分或完全函数依赖

在 $R(U)$ 中, 若 $X \rightarrow Y$ 并且对于 X 的任何真子集 X' 都有 $X' \not\rightarrow Y$, 则称 Y 完全函数依赖于 X , 记为: $X \xrightarrow{f} Y$, 否则称 Y 部分函数依赖于 X , 记为: $X \xrightarrow{p} Y$

示例: $U = \{\text{学号}, \text{姓名}, \text{年龄}, \text{班号}, \text{班长}, \text{课号}, \text{成绩}\}$

□ $\{\text{学号}, \text{课号}\} \xrightarrow{f} U$

□ $\{\text{学号}, \text{课号}\} \xrightarrow{p} \text{姓名}$

□ $\{\text{学号}, \text{课号}\} \xrightarrow{f} \text{成绩}$

部分依赖
存在着非
受控冗余!

学生				
学号	姓名	课程号	课程名	成绩
98030101	张三	001	数据库	92
98030101	张三	002	计算机原理	85
98030101	张三	003	高等数学	88
98040202	李四	002	计算机原理	90
98040202	李四	003	高等数学	80
98040202	李四	001	数据库	55
98040203	王五	003	高等数学	56
98030102	周六	001	数据库	54
98030102	周六	002	计算机原理	85
98030102	周六	003	高等数学	48

练习：分析下列模式的完全或部分函数依赖

- 学生(学号, 姓名, 班级, 课号, 课程名, 成绩, 教师, 教师职务)
- 员工(员工码, 姓名, 出生日期, 联系电话, 最后学历, 毕业学校, 培训日期, 培训内容)
- 图书(书号, 书名, 出版日期, 出版社, 书架号, 房间号)
- 客户(客户号, 客户名称, 类别, 联系电话, 产品编码, 产品名称, 数量, 要货日期)
- 学生(学号, 姓名, 系号, 系主任)

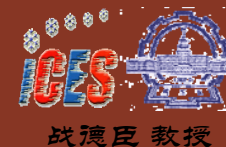
练习：分析下列模式的完全或部分函数依赖

- 学生(学号, 姓名, 班级, 课号, 课程名, 成绩, 教师, 教师职务)
 - {学号, 课号} \xrightarrow{f} U ; 但 {学号, 课号} \xrightarrow{p} 姓名
 - {学号, 课号} \xrightarrow{p} 课程名
- 员工(员工码, 姓名, 出生日期, 联系电话, 最后学历, 毕业学校, 培训日期, 培训内容)
 - {员工码, 培训日期} \xrightarrow{f} U ;
 - {员工码, 培训日期} \xrightarrow{p} { 姓名, 出生日期 } ;
- 图书(书号, 书名, 出版日期, 出版社, 书架号, 房间号)
- 客户(客户号, 客户名称, 类别, 联系电话, 产品编码, 产品名称, 数量, 要货日期)
- 学生(学号, 姓名, 系号, 系主任)

其他几个,
同学自己
分析...

完全函数依赖与传递函数依赖

(3)传递函数依赖的定义



[Definition]传递函数依赖

在 $R(U)$ 中, 若 $X \rightarrow Y$, $Y \rightarrow Z$ 且 $Y \not\subset X$, $Z \not\subset Y$, $Z \not\subset X$, $Y \not\rightarrow X$, 则称 Z 传递函数依赖于 X 。

示例： $U = \{\text{学号}, \text{姓名}, \text{年龄}, \text{班号}, \text{班长}, \text{课号}, \text{成绩}\}$

□ 学号 \rightarrow 班号 ; 班号 \rightarrow 班长

□ 学号 \rightarrow 班长

注：“班长”是传递依赖于“学号”的。

示例： 学生(学号, 姓名, 系号, 系主任)

□ 学号 \rightarrow 系号 ; 系号 \rightarrow 系主任

□ 学号 \rightarrow 系主任

注：“系主任”是传递依赖于“学号”的。

学号	姓名	班级	班主任	班主任职称
2003510101	张三	035101	张林	讲师
2003510102	李四	035101	张林	讲师
2003510103	王五	035101	张林	讲师
2003510104	李六	035101	张林	讲师
2003510105	张四	035101	张林	讲师
2003510106	张五	035101	张林	讲师
2003510107	张小三	035101	张林	讲师
2003510108	张小四	035101	张林	讲师
2003510109	李小三	035101	张林	讲师
2003510110	李小四	035101	张林	讲师
2003520201	周三	035202	郑东	副教授
2003520202	赵四	035202	郑东	副教授
2003520203	赵五	035202	郑东	副教授
2003520204	赵六	035202	郑东	副教授
2003520205	钱四	035202	郑东	副教授
2003520206	强五	035202	郑东	副教授
2003520207	梁小三	035202	郑东	副教授
2003520208	梁小四	035202	郑东	副教授
2003520209	王小三	035202	郑东	副教授
2003520210	王小四	035202	郑东	副教授

传递依赖
存在着非
受控冗余!

完全函数依赖与传递函数依赖

(4)传递函数依赖的示例



练习：分析下列模式的传递函数依赖

- 商店(商店, 商品, 商品经营部, 经营部经理)
- 学生(学号, 姓名, 班级, 班主任, 课号, 课程名, 成绩, 教师, 教师职务)
- 员工(员工码, 姓名, 部门, 部门经理)
- 图书(书号, 书名, 出版日期, 出版社, 书架号, 房间号, 管理员)
- 客户(客户号, 客户名称, 类别, 联系电话, 产品编码, 产品名称, 数量, 要货日期)

完全函数依赖与传递函数依赖

(4)传递函数依赖的示例



练习：分析下列模式的传递函数依赖

- 商店(商店, 商品, 商品经营部, 经营部经理)
 - {商店, 商品} → 商品经营部 ; {商店, 商品经营部} → 经营部经理
 - {商店, 商品} → 经营部经理
- 学生(学号, 姓名, 班级, 班主任, 课号, 课程名, 成绩, 教师, 教师职务)
 - 学号 → 班级 ; 班级 → 班主任 ; {学号, 课号} → 教师 ; 教师 → 教师职务
 - 学号 → 班主任 ; {学号, 课号} → 教师职务
- 员工(员工码, 姓名, 部门, 部门经理)
 - 员工码 → 部门 ; 部门 → 部门经理
 - 员工码 → 部门经理
- 图书(书号, 书名, 出版日期, 出版社, 书架号, 房间号, 管理员)
- 客户(客户号, 客户名称, 类别, 联系电话, 产品编码, 产品名称, 数量, 要货日期)

其他两个,
同学自己
分析...

函数依赖相关的几个重要概念

战德臣

哈尔滨工业大学 教授.博士生导师
黑龙江省教学名师
教育部大学计算机课程教学指导委员会委员

Research Center on Intelligent
Computing for Enterprises & Services,
Harbin Institute of Technology

[Definition] 候选键

设K为R(U)中的属性或属性组合，若 $K \xrightarrow{f} U$ ，则称K为R(U)上的**候选键** (Candidate Key)。

说明：

- (1) 可任选一候选键作为R的**主键** (Primary Key)；
- (2) 包含在任一候选键中的属性称**主属性** (Prime Attribute)，其他属性称**非主属性**；
- (3) 若K是R的一个候选键， $S \supset K$ ，则称S为R的一个**超键** (Super Key)。

示例： $U = \{\text{学号}, \text{姓名}, \text{年龄}, \text{班号}, \text{班长}, \text{课号}, \text{成绩}\}$

□ $\{\text{学号}, \text{课号}\} \xrightarrow{f} U$

函数依赖相关的几个重要概念

(2)候选键的示例

练习：找候选键与非主属性

- 学生(学号, 年龄, 家庭住址, 课程号, 成绩, 教师, 教师职务)

候选键是??, 非主属性是??

- 邮编(城市名, 街道名, 邮政编码)

候选键是??, 非主属性是??

- 商店(商店, 商品, 商品经营部, 商品经营部经理)

候选键是??, 非主属性是??

- 学生(学号, 姓名, 所属系别, 系主任)

候选键是??, 非主属性是??

找候选键需
要先找出函
数依赖集合

[Definition] 外来键

若 $R(U)$ 中的属性或属性组合 X 并非 R 的候选键，但 X 却是另一关系的候选键，则称 X 为 R 的外来键(Foreign Key)，简称外键。

示例： $R = \{ \text{合同号}, \text{合同名}, \text{签订日期}, \text{供应商名} \}$

$S = \{ \text{供应商名}, \text{地址}, \text{执照号}, \text{法人代表} \}$

□ 合同号 \xrightarrow{f} $\{ \text{合同号}, \text{合同名}, \text{签订日期}, \text{供应商名} \}$

□ 供应商名 \xrightarrow{f} $\{ \text{供应商名}, \text{地址}, \text{执照号}, \text{法人代表} \}$

理解外键需
要先理解候
选键

一个关系的外
键是另一个关
系的候选键

函数依赖相关的几个重要概念

(4)逻辑蕴涵的定义



[Definition] 逻辑蕴涵

设 F 是关系模式 $R(U)$ 中的一个函数依赖集合, X, Y 是 R 的属性子集, 如果从 F 中的函数依赖能够推导出 $X \rightarrow Y$, 则称 F 逻辑蕴涵 $X \rightarrow Y$, 或称 $X \rightarrow Y$ 是 F 的逻辑蕴涵。记作 $F \models X \rightarrow Y$ 。

说明:

- 设 F 是关系模式 $R(U)$ 的函数依赖集, $X \rightarrow Y$ 是一个函数依赖, 若对 R 中的每个满足 F 的关系 r , 能够用形式逻辑推理的方法推出 r 也满足 $X \rightarrow Y$, 则称 $F \models X \rightarrow Y$ 。
- 若满足 F 的每个关系均满足 $X \rightarrow Y$, 则说 F 逻辑蕴涵 $X \rightarrow Y$;

函数依赖有显性给出的, 也有隐含的

怎样推导, 在后面讲授

函数依赖相关的几个重要概念

(5)闭包的定义



[Definition]闭包

被F逻辑蕴涵的所有函数依赖集合称为F的闭包(Closure), 记作 F^+ 。

说明：

□若 $F^+ = F$, 则说F是一个全函数依赖族(函数依赖完备集)。

示例：设 $R=ABC$, $F=\{ A \rightarrow B, B \rightarrow C \}$, 则 F^+ 的组成如下, 即由如下形式的 $X \rightarrow Y$ 构成：

- ✓(1) X包含A, 而Y任意, 如 $ABC \rightarrow AB, A \rightarrow C, AB \rightarrow BC, \dots$
- ✓(2) X包含B但不含A且Y不含A, 如 $BC \rightarrow B, B \rightarrow C, B \rightarrow \phi, \dots$
- ✓(3) $X \rightarrow Y$ 是 $C \rightarrow C$ 或 $C \rightarrow \phi$

小集合
大闭包

包含了平
凡的函数
依赖

怎样求隐含
的函数依赖
呢?

关于函数依赖的公理和定理

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

关于函数依赖的公理和定理

(1)函数依赖的Armstrong公理



[Armstrong's Axioms A1~A3]

设 $R(U)$ 是属性集 $U=\{A_1, A_2, \dots, A_n\}$ 上的一个关系模式， F 为 $R(U)$ 的一组函数依赖，记为 $R(U, F)$ ，则有如下规则成立：

□[A1]自反律(Reflexivity rule)：若 $Y \subseteq X \subseteq U$ ，则 $X \rightarrow Y$ 被 F 逻辑蕴涵。

□[A2]增广律(Augmentation rule)：若 $X \rightarrow Y \in F$ ，且 $Z \subseteq U$ ，则 $XZ \rightarrow YZ$ 被 F 逻辑蕴涵。

□[A3]传递律(Transitivity rule)：若 $X \rightarrow Y \in F$ ，且 $Y \rightarrow Z$ ，则 $X \rightarrow Z$ 被 F 逻辑蕴涵。

公理的作用是由已知的函数依赖推导出隐含的函数依赖

[引理1]Armstrong's Axiom规则A1, A2, A3是有效的(正确的)。

证明：设 r 是 $R(U, F)$ 的任一关系，有任意两个元组 $t, s \in r$

<**A1正确性证明**>：因为若 $t[X]=s[X]$, $Y \subseteq X$, 则可推出 $t[Y]=s[Y]$, 所以： $X \rightarrow Y$ 。

<**A2正确性证明**>：因为若 $t[XZ] = s[XZ]$, 则应有 $t[X]=s[X]$, $t[Z] = s[Z]$;若 $t[YZ] = s[YZ]$, 则应有 $t[Y] = s[Y]$, $t[Z] = s[Z]$;由 $X \rightarrow Y$ 可知若 $t[X] = s[X]$, 则一定有 $t[Y] = s[Y]$; 因此, 若 $t[XZ] = s[XZ]$, 则一定可推出 $t[YZ] = s[YZ]$, 所以 $XZ \rightarrow YZ$ 。

<**A3正确性证明**>：因为：由 $X \rightarrow Y$ 可知若 $t[X] = s[X]$, 则一定有 $t[Y] = s[Y]$; 由 $Y \rightarrow Z$ 可知若 $t[Y] = s[Y]$, 则一定有 $t[Z] = s[Z]$; 因此, 若 $t[X] = s[X]$, 则一定可推出 $t[Z] = s[Z]$, 所以 $X \rightarrow Z$ 。 证毕。

公理的正确性
需要依据
定义来证明

[引理2]由Armstrong 's Axiom可推出如下结论：

- (a) **合并律**(Union Rule)：若 $X \rightarrow Y$ 且 $X \rightarrow Z$, 则 $X \rightarrow YZ$ 。
- (b) **伪传递律**(Pseudo Transitivity)：若 $X \rightarrow Y$ 且 $WY \rightarrow Z$, 则 $XW \rightarrow Z$ 。
- (c) **分解律**(Decomposition Rule)：若 $X \rightarrow Y$ 且 $Z \subseteq Y$, 则 $X \rightarrow Z$ 。

证明：(a)合并律的正确性证明：由 $X \rightarrow Y$ 和增广律，可以推出 $X \rightarrow XY$ (注：两边都增加一个X); 由 $X \rightarrow Z$ 和增广律，可以推出 $XY \rightarrow YZ$ (注：两边都增加一个Y);再由传递律，可得： $X \rightarrow YZ$ 。

(b)伪传递律的正确性证明：由 $X \rightarrow Y$ 和增广律，可以推出 $WX \rightarrow WY$ (注：两边都增加一个W);又由 $WY \rightarrow Z$ 和传递律，可以推出 $XW \rightarrow Z$ 。

(c)分解律的正确性证明：由 $Z \subseteq Y$ 和自反律，可以推出， $Y \rightarrow Z$ ；再由 $X \rightarrow Y$ 及传递律，可以推出 $X \rightarrow Z$ 。
证毕。

[引理3]如果 A_1, A_2, \dots, A_n 是属性，则 $X \rightarrow A_1, A_2, \dots, A_n$ 当且仅当对每个 A_i 有 $X \rightarrow A_i (1 \leq i \leq n)$ 。

证明略。

关于属性组合的
函数依赖与单一
属性的函数依赖
有什么关系？

定理的正确性
可依据公理和
其他已证明之
定理来证明

关于函数依赖的公理和定理

(4)属性闭包



[Definition] 属性(集)闭包

对 $R(U, F)$, $X \subseteq U$, $U = \{A_1, A_2, \dots, A_n\}$, 令 :

$$X^+_F = \{ A_i \mid \text{用Armstrong Axiom A1,A2,A3可从F导出} X \rightarrow A_i \}$$

称 X^+_F 为 X 关于 F 的属性(集)闭包。

注：显然 $X \subseteq X^+_F$ 。

[引理4] $X \rightarrow Y$ 可从 F 由Armstrong Axiom导出，当且仅当 $Y \subseteq X^+_F$ 。

证明：<首先证明必要性：若 $X \rightarrow Y$ ，则定有 $Y \subseteq X^+_F$ >

因为 $X \rightarrow Y$ ，由分解律，若 $A_i \in Y$ ，则 $X \rightarrow A_i$ ($i = 1, 2, \dots, n$)，所以 $Y \subseteq X^+_F$ 。

<再证明充分性：若 $Y \subseteq X^+_F$ ，则定有 $X \rightarrow Y$ >

因为 $Y \subseteq X^+_F$ ，按属性闭包定义，对 $\forall A_i \in Y$ ，有 $X \rightarrow A_i$ ($i = 1, 2, \dots, n$)；再由合并律，可以推出 $X \rightarrow Y$ 。证毕。

如何证明一个函数依赖被蕴涵？

关于函数依赖的公理和定理

(5)Armstrong公理的完备性及其证明

[定理] Armstrong Axiom A1,A2,A3是有效的和完备的

公理的有效性
是指通过公理
推出的结论是
正确的

公理的完备性是指被F
逻辑蕴涵的所有函数依
赖都能由公理A1,A2,A3
在F基础上推出

[定理]: Armstrong Axiom A1, A2, A3 是有效的和完备的。

证明: 公理的有效性是指通过公理推出的结论是正确的。而公理的完备性是指被 F 逻辑蕴涵的所有函数依赖都能由公理 A1, A2, A3 在 F 基础上推出, 也就是说, 若基于 F 能由公理 A1, A2, A3 导出 $X \rightarrow Y$, 则必有 F 逻辑蕴涵 $X \rightarrow Y$; 否则 F 必不逻辑蕴涵 $X \rightarrow Y$ 。公理的有效性已在引理 1 中给出。下面只需证明完备性。

证明完备性可以从能否找出一个满足下面要求的关系 r 来进行: r 使 F 的所有函数依赖都成立(即 r 上满足 F)。由于假设 $X \rightarrow Y$ 不能从 F 用公理推出, 则 r 上的 $X \rightarrow Y$ 肯定不成立。如果找到 r , 则 F 必不逻辑蕴涵 $X \rightarrow Y$ 。

设 F 为 $R(U)$ 上的一组函数依赖, $X \rightarrow Y$ 不能用公理由 F 导出, 考虑如下的一个具体关系 r :

r :	X^+ 的属性	其他属性
	1 1 1 ... 1	1 1 1 ... 1
	1 1 1 ... 1	0 0 0 ... 0

这是一个特
殊的关系,
两行两列

(1)先证明 r 满足 F 中的全部函数依赖。

用反证法证明, 假设存在一个函数依赖 $V \rightarrow W \in F$, 但不被 r 满足。

再假设有 $V \subseteq X^+$ (如果 $V \not\subseteq X^+$, 则 r 中没有两个相等的 V , 则 $V \rightarrow W$ 将能够被 r 满足)

则必有 $W \not\subseteq X^+$ (否则的话, 如果 $W \subseteq X^+$, 因为 $V \subseteq X^+$, r 中相应于 X^+ 的属性值都为 1, 则 $V \rightarrow W$ 也将能够被 r 满足)

因此假设 A 是 W 中不属于 X^+ 的属性, 由自反律, $W \rightarrow A$

因为 $V \subseteq X^+$, 由引理 4, $X \rightarrow V$ 能由公理导出。

又因为 $V \rightarrow W \in F$, 由传递律有 $X \rightarrow W$, 所以再由传递律有 $X \rightarrow A$

由引理 4, 可推出 $A \subseteq X^+$, 这与 A 的假设矛盾, 故在 F 中的任何 $V \rightarrow W$ 一定被 r 满足。

(2)再证明 $X \rightarrow Y$ 不被 r 满足。

也可用反证法证明, 假设 r 满足 $X \rightarrow Y$, 显然有 $X \subseteq X^+$, $Y \subseteq X^+$, 由引理 4, $X \rightarrow Y$ 将能由公理导出, 这与假设矛盾。所以 $X \rightarrow Y$ 不被 r 满足。

所以可推出: 只要 $X \rightarrow Y$ 不能由 F 用 Armstrong Axiom 导出, 则 F 必定不逻辑蕴涵 $X \rightarrow Y$ 。即 Armstrong Axiom A1, A2, A3 是完备的。证毕。

函数依赖集的最小覆盖

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

函数依赖集的最小覆盖

(1)覆盖的概念



[Definition] **覆盖(Cover)**

对 $R(U)$ 上的两个函数依赖集合 F 、 G ，如果 $F^+ = G^+$ ，则称 F 和 G 是等价的，也称 F 覆盖 G 或者 G 覆盖 F 。

[引理5]: $F^+ = G^+ \Leftrightarrow F \subseteq G^+ \wedge G \subseteq F^+$

证明： \Rightarrow (必要性)：因为： $F \subseteq F^+$ 以及条件 $F^+ = G^+$ 所以 $F \subseteq G^+$ ，同理可证 $G \subseteq F^+$

\Rightarrow (充分性)：由条件： $F \subseteq G^+$ ，可推出 $X_F^+ \subseteq X_{G^+}^+$ 也就是说，对任一 $X \rightarrow Y \in F^+$ ，有 $Y \subseteq X_F^+$ ，从而 $Y \subseteq X_{G^+}^+$ ；进一步可推出， $X \rightarrow Y \in (G^+)^+ = G^+$ ，所以 $F^+ \subseteq G^+$ ；同理可证， $G^+ \subseteq F^+$ 。
所以，最终结论 $F^+ = G^+$ 成立。

为验证 F 、 G 是否等价，需要对任一 $X \rightarrow Y \in F^+$ 检验是否 $X \rightarrow Y \in G^+$ ，由引理4可将这种检验转为是否 $Y \subseteq X_{G^+}^+$ ，由此可将求 F^+ 这种不容易完成的事转为容易完成的事，即求属性闭包。

函数依赖集的最小覆盖

(2)属性闭包的计算算法



[Algorithm] 计算一属性集关于一组函数依赖的属性闭包。

Input : 有限属性集合U, U上的函数依赖集合F, 及U的子集X

Output : X关于F的**属性闭包** X^+ , 记为 X_F^+ 。

Method : 按下列规则递归计算属性序列 $X^{(0)}, X^{(1)}, \dots$

1. 令 $X^{(0)}=X, i=0$
2. $B = \{A \mid (\exists V)(\exists W)(V \rightarrow W \in F \wedge V \subseteq X^{(i)} \wedge A \subseteq W)\}$
3. $X^{(i+1)} = B \cup X^{(i)}$
4. If $X^{(i+1)} \neq X^{(i)}$ then $i=i+1$; goto 2.
5. $X_F^+ = X^{(i)}$, 算法终止。

示例 : 已知 $R(U, F), U=\{A, B, C, D, E\}, F=\{AB \rightarrow C, B \rightarrow D, C \rightarrow E, EC \rightarrow B, AC \rightarrow B\}$ 。 求 : $(AB)_F^+$

解 : (1) $X^{(0)} = \{A, B\}$

(2) 由 $AB \rightarrow C, B \rightarrow D$: $X^{(1)} = X^{(0)} \cup \{C, D\} = \{A, B, C, D\}$

(3) 由 $C \rightarrow E, AC \rightarrow B$: $X^{(2)} = X^{(1)} \cup \{E\} = \{A, B, C, D, E\}$

(4) 由 $EC \rightarrow B$: $X^{(3)} = X^{(2)} \cup \phi = \{A, B, C, D, E\}$

(5) 因为 $X^{(3)} = X^{(2)}$, 所以 $(AB)_F^+ = \{A, B, C, D, E\}$ 。

函数依赖集的最小覆盖 (2)属性闭包的计算算法

**[定理] Algorithm 正确地
计算了 X^+_F 。**

[定理]: Algorithm 正确地计算了 X^+ 。

证明: (1)先证若属性 A 在 $X^{(0)}$ 中, 则 A 必在 X^+_F 中。可用归纳法证明。

Basis: $j=0$, 若 $A \in X^{(0)}$, 即是 $A \in X$, 自然就有 $X \rightarrow A$, $A \in X^+_F$

Induction: 设 $j>0$, $X^{(j-1)}$ 中的属性均在 X^+ 中

依据算法规则假设 $A \in Z$, $Y \rightarrow Z \in F$, $Y \subseteq X^{(j-1)}$, 则有 $A \in X^{(j)}$

因为 $Y \subseteq X^{(j-1)}$ 按归纳假设 $Y \subseteq X^+$, 由引理 4, 可推出 $X \rightarrow Y$

又因为 $Y \rightarrow Z$, 由传递律知 $X \rightarrow Z$, 再由自反律知 $Z \rightarrow A$, 所以 $X \rightarrow A$, A 在 X^+_F 中。

(2)再证若 $A \in X^+$, 则 $A \in X^{(j)}$

在下面的证明中, 我们将利用<事实*>: 对 M_1, M_2 , 若 $M_1 \subseteq M_2$, 则对 $\forall j, M_1^{(j)} \subseteq M_2^{(j)}$ 。

我们可用归纳法证明断言(2), 即用对由公理导出 $X \rightarrow A$ 的步数归纳证明 A 在某个 $X^{(j)}$ 中。

Basis: 一步导出, 此时 $X \rightarrow Y$ 要么按自反律得到, 要么本身在 F 中, 对前者 $Y \subseteq X^{(0)}$, 对后者 $Y \subseteq X^{(1)}$ 。

Induction: 假设上述断言(2)对少于 p 步的导出是真的, 而 $X \rightarrow Y$ 是经过 p 步导出的。

首先假设 $X \rightarrow Y$ 是由 $X \rightarrow Z$, $Z \rightarrow Y$ 依据传递律导出的, 因为 $X \rightarrow Z$, $Z \rightarrow Y$ 的导出少于 p 步, 按归纳假设, $\exists j$ 使 $Z \subseteq X^{(j)}$ 。若以 Z 代替 X 应用算法 1, 则由归纳假设知, $\exists k$ 使 $Y \subseteq Z^{(k)}$ 。若令 $M_1=Z$, $M_2=X^{(j)}$, 则由前述<事实*>, 知 $Y \subseteq X^{(j+k)}$ (因为 $Y \subseteq Z^{(k)} = M_1^{(k)} \subseteq M_2^{(k)} = (X^{(j)})^{(k)} = X^{(j+k)}$)。

再假设 $X \rightarrow Y$ 或在 F 中或按自反律导出, 则有, 若 $X \rightarrow Y \in F$, 由归纳假设知, $\exists j, X \subseteq X^{(j)}$, 由此可推出 $Y \subseteq X^{(j+1)}$ 。若 $X \rightarrow Y$ 由自反律导出, 由归纳假设知, $\exists j, X \subseteq X^{(j)}$, 由此可推出 $Y \subseteq X^{(j)}$ 。

再假设 $X \rightarrow Y$ 由前面的 $V \rightarrow W$ 利用增广律导出, 设 $VZ=X$, $WZ=Y$, 因为 $V \rightarrow W$ 的导出少于 p 步, 按归纳假设知, $\exists j$ 使 $W \subseteq V^{(j)}$, 对 $X=VZ$ 应用算法 1, 则利用<事实*>, 令 $M_1=V$, $M_2=X$, 有 $W \subseteq X^{(j)}$ (因为 $W \subseteq V^{(j)} \subseteq (VZ)^{(j)} = X^{(j)}$)。又因为 $Z \subseteq X \subseteq X^{(j)}$, 所以 $Y=WZ \subseteq X^{(j)}$, 归纳完毕。

又由引理 4, 若 $A \in X^+$, 则 $X \rightarrow A$ 能由公理基于 F 推出, 所以必有 $\exists j, A \in X^{(j)}$ 。证毕。

同学可自行
阅读此定理
的证明...

[引理6] 每个函数依赖集F可被一个其右端至多有一个属性的函数依赖之集G覆盖。

证明：依据引理6， $G = \{ X \rightarrow A \mid X \rightarrow Y \in F \wedge A \in Y \}$

要证 F^+ 与 G^+ 等价，由引理5，只需证 $F \subseteq G^+ \wedge G \subseteq F^+$

因为 $X \rightarrow Y \in F$ ， $A \in Y$ 可以由Armstrong公理的自反律可推出 $X \rightarrow A$ ，所以 $G \subseteq F^+$

又因为若 $Y = A_1 A_2 \dots A_n$ ，则G中的 $X \rightarrow A_1$ ， $X \rightarrow A_2$ ，...， $X \rightarrow A_n$ ，由Armstrong公理的合并律可以推出 $X \rightarrow Y$ ，所以 $F \subseteq G^+$ 。最终依据引理5，可导出 $F^+ = G^+$ 。证毕。

函数依赖集的最小覆盖

(4)最小覆盖



➤[Definition] **最小覆盖**

若F满足以下条件，则称F为最小覆盖(minimal Cover)或最小依赖集(minimal set of Functional Dependency)：

- 1) F中每个函数依赖的右部是单个属性；
- 2) 对任何 $X \rightarrow A \in F$, 有 $F - \{ X \rightarrow A \}$ 不等价于F；
- 3) 对任何 $X \rightarrow A \in F$, $Z \subset X$, $(F - \{ X \rightarrow A \}) \cup \{ Z \rightarrow A \}$ 不等价于F。

➤[定理]：每个函数依赖集F都有等价的最小覆盖F'。

为什么
要求最
小覆盖？

怎样求
最小覆
盖？

证明兼算法：由引理6可假设F等价于一个依赖集G, G中没有一个右端多于一个的属性。

再考虑G中每一依赖 $X \rightarrow Y$, 取一次序进行检查：若 $G - \{ X \rightarrow Y \}$ 等价于G, 则由G中删去 $X \rightarrow Y$ 。由此可得新的依赖集G'。

对G'中的每一依赖，取一次序进行检查：考虑其左端属性，若消去左端某一属性后，还保持等价性，则消去它，由此可得新的依赖集F'。

显然F'等价于F, 而F'就是一个最小覆盖。

定理证毕。

回顾本讲学了什么？

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

回顾本讲学习了什么？

