

第7章 關聯式資料庫的正規化

1

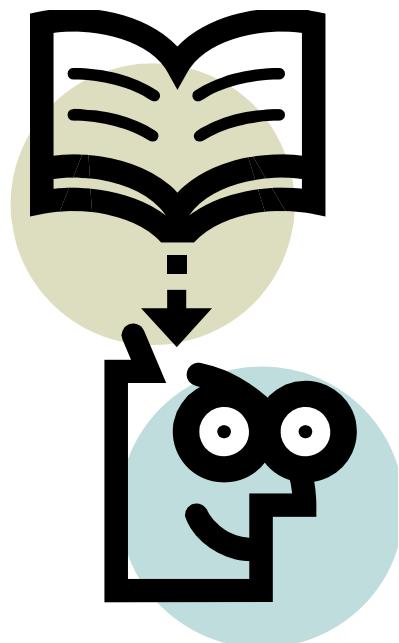
- 7-1 正規化的基礎
- 7-2 功能相依
- 7-3 第一階到第三階正規化型式
- 7-4 多重值相依與第四階正規化型式
- 7-5 合併相依與第五階正規化型式



7-1 正規化的基礎

2

- 7-1-1 關聯表為什麼需要正規化
- 7-1-2 正規化的型式



7-1 正規化的基礎-說明

3

- 「正規化」（Normalization）是一種標準處理程序來決定關聯表應該擁有哪些屬性，其目的是建立「良好結構關聯表」（Well-structured Relation）。
- 良好結構關聯表是一種沒有重複資料的關聯表，而且在新增、刪除或更新資料時，不會造成錯誤或資料不一致的異常情況。

7-1 正規化的基礎-目的

4

- **去除重複性（ Eliminating Redundancy ）**：就是建立沒有重複資料的關聯表，因為重複資料不只浪費資料庫的儲存空間，而且會產生資料維護上的問題。
- **去除不一致的相依性（ Eliminating Inconsistent Dependency ）**：資料相依是指關聯表中的屬性之間擁有關係，如果關聯表擁有不一致的資料相依，這些屬性就會在新增、刪除或更新資料時，造成異常現象。

7-1-1 關聯表為什麼需要正規化

5

- 關聯式資料庫正規化的目的以一句話來說，就是**避免資料重複**。**WHY?** 因為將不應該放在同一個關聯表的屬性**誤放或硬置**於同一個關聯表，形成**非自然結合**。
- 關聯表擁有非自然結合的屬性，就會造成資料重複的問題，稱為「**非正規化型式**」（**Unnormalized Form**）關聯表。

「非正規化型式」關聯表

6

- 存在**非自然結合的屬性**, e.g., {name, birthday}, {title, credit}

Student_Course_Classes

sid	name	birthday	c_no	title	credits	time	room
S001	陳會安	1967/9/3	CS101	計算機概論	4	12:00pm	180-M
S003	張三丰	1982/3/3	CS121	離散數學	4	8:00am	221-S
S001	陳會安	1967/9/3	CS203	程式語言	3	10:00am	221-S
S002	江小魚	1978/2/2	CS203	程式語言	3	14:00pm	327-S
S001	陳會安	1967/9/3	CS222	資料庫管理系統	3	13:00pm	100-M
S002	江小魚	1978/2/2	CS222	資料庫管理系統	3	13:00pm	100-M
S004	李四方	1981/4/4	CS222	資料庫管理系統	3	13:00pm	100-M
S003	張三丰	1982/3/3	CS213	物件導向程式設計	2	9:00am	622-G
S001	陳會安	1967/9/3	CS213	物件導向程式設計	2	12:00pm	500-K

「非正規化型式」關聯表- 更新異常(Update Anomaly)

7

- 在Student_Course_Classes關聯表更新指定的屬性資料。因為資料重複儲存在多筆值組（記錄），需要同時更新多筆值組的資料。

Student_Course_Classes

sid	name	birthday	c_no	title	credits	time	room
S001	陳會安	1965/9/3	CS101	計算機概論	4	12:00pm	180-M
S001	陳會安	1965/9/3	CS203	程式語言	3	10:00am	221-S
S001	陳會安	1965/9/3	CS222	資料庫管理系統	3	13:00pm	100-M
S001	陳會安	1965/9/3	CS213	物件導向程式設計	2	12:00pm	500-K

「非正規化型式」關聯表- 刪除異常(Deletion Anomaly)

8

- 在Student_Course_Classes關聯表刪除值組，可能一併刪除了合法資料。
- 例如：CS101和CS121課程都有一筆值組(第1, 2筆值組)，故若刪除掉這2筆值組，則將會遺失了CS101和CS121課程的資訊。
(只因沒有人選修!)
- 如下頁

「非正規化型式」關聯表- 刪除異常(續)

9

- 刪除值組，可能一併刪除了合法資料

Student_Course_Classes

sid	name	birthday	c_no	title	credits	time	room
S001	陳會安	1967/9/3	CS101	計算機概論	4	12:00pm	180-M
S003	張三丰	1982/3/3	CS121	離散數學	4	8:00am	221-S
S001	陳會安	1967/9/3	CS203	程式語言	3	10:00am	221-S
S002	江小魚	1978/2/2	CS203	程式語言	3	14:00pm	327-S

.....

S003	張三丰	1982/3/3	CS213	物件導向程式設計	2	9:00am	622-G
S001	陳會安	1967/9/3	CS213	物件導向程式設計	2	12:00pm	500-K

「非正規化型式」關聯表- 新增異常(Insertion Anomaly)

10

- 在Student_Course_Classes關聯表新增值組和刪除相反，可能只新增部分值組的資料。例如：新增一筆值組，課程編號是CS240，課程名稱為”資料結構”，學分是3，但是sid出現Null，如下：

Student_Course_Classes

sid	name	birthday	c_no	title	credits	time	room
S001	陳會安	1967/9/3	CS101	計算機概論	4	12:00pm	180-M
S003	張三丰	1982/3/3	CS121	離散數學	4	8:00am	221-S
S001	陳會安	1967/9/3	CS203	程式語言	3	10:00am	221-S
S002	江小魚	1978/2/2	CS203	程式語言	3	14:00pm	327-S

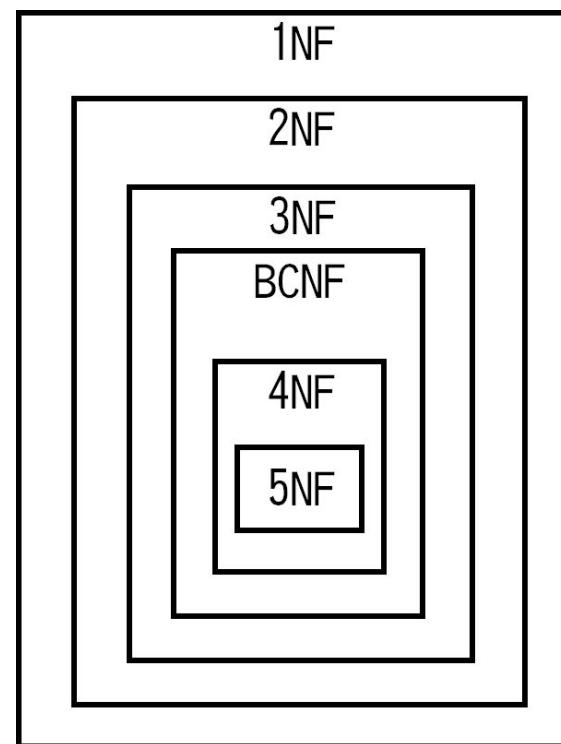
.....

S003	張三丰	1982/3/3	CS213	物件導向程式設計	2	9:00am	622-G
S001	陳會安	1967/9/3	CS213	物件導向程式設計	2	12:00pm	500-K
--	--	--	CS240	資料結構	3	3:00pm	230-A

7-1-2 正規化的型式-階段

11

- 正規化的處理過程一共分為五個階段，每一階正規化型式是在處理不同的屬性間資料相依（Dependency）的問題，如下圖所示：



7-1-2 正規化的型式-前三階

12

- **第一階正規化型式 (First Normal Form, 1NF) :**
在關聯表刪除多重值屬性，讓關聯表只擁有單元值屬性。
- **第二階正規化型式 (Second Normal Form, 2NF) :**
滿足1NF且關聯表沒有「部分相依」 (Partial Dependency) 。
- **第三階正規化型式 (Third Normal Form, 3NF) :**
滿足2NF且關聯表沒有「遞移相依」 (Transitive Dependency) 。

7-1-2 正規化的型式-BCNF、4、5階

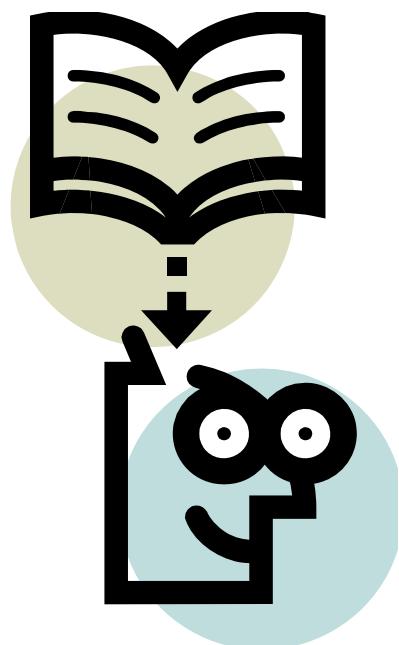
13

- **Boyce-Codd正規化型式（Boyce-Codd Normal Form, BCNF）**：屬於一種廣意的第三階正規化型式，如果關聯表擁有多個複合候選鍵，刪除候選鍵間的功能相依。
- **第四階正規化型式（Fourth Normal Form, 4NF）**：滿足BCNF且沒有多重值相依。
- **第五階正規化型式（Fifth Normal Form, 5NF）**：滿足4NF且沒有合併相依。

7-2 功能相依

14

- 7-2-1 功能相依的定義
- 7-2-2 功能相依的種類
- 7-2-3 找出所有的功能相依
- 7-2-4 阿姆斯壯推論規則
- 7-2-5 功能相依的推論範例



7-2 功能相依-說明

15

- 功能相依 (Functional Dependency , 簡稱FD) 是描述關聯表中屬性間的相依關係，這是關聯式資料庫正規化的基礎。
- 簡單的說，關聯表正規化的第一步，就是在關聯表找出所有屬性間的功能相依。

7-2 功能相依-目的

16

- 當在關聯表找出所有屬性間的功能相依後，就可以幫助我們在關聯表找出：
 - **重複資料**：不該屬於此關聯表的屬性。
 - **候選鍵和主鍵**：如果關聯表的所有屬性都功能相依於一個或一組屬性，這個屬性就是候選鍵（Candidate Key）或主鍵（Primary Key）。

7-2-1 功能相依的定義-定義

17

定義7.1：功能相依（Functional Dependency）以關聯表R為例，若關聯表R擁有屬性A與B（A與B可以是複合屬性），則若屬性B功能相依（Functional dependent）於A，或A功能決定（Functional determines）B，可表示為：

$R.A \rightarrow R.B$

或簡寫為：

$A \rightarrow B$

若關聯表R擁有 $A \rightarrow B$ 的功能相依，則在關聯表R的每一對 t_1 和 t_2 值組且 $t_1(A) = t_2(A)$ ，可以得到： $t_1(B) = t_2(B)$ 。

7-2-1 功能相依的定義-說明

18

- 關聯表中如果有2個值組的屬性A值是相同，我們就可以知道屬性B的值也一定相同。換句話說，如果知道屬性A的值，就可以知道屬性B的值，如：

$A_1 \rightarrow B_1$

$A_1 \rightarrow \{ B_1, B_2, \dots, B_n \}$

$\{ A_1, A_2, \dots, A_n \} \rightarrow B_1$

$\{ A_1, A_2, \dots, A_n \} \rightarrow \{ B_1, B_2, \dots, B_n \}$

- 功能相依屬於一種單方向的關聯， $A \rightarrow B$ 不表示 $B \rightarrow A$ 。

- 功能相依是一種屬性間的關係

7-2-1 功能相依的定義-範例

19

- 在Students關聯表的一些功能相依範例，如下：

$\text{sid} \rightarrow \text{birthday}$

$\text{sid} \rightarrow \text{GPA}$

$\text{sid} \rightarrow \text{tel}$

$\text{sid} \rightarrow \{ \text{name}, \text{birthday} \}$

$\text{sid} \rightarrow \{ \text{name}, \text{birthday}, \text{GPA} \}$

$\text{sid} \rightarrow \{ \text{name}, \text{birthday}, \text{GPA}, \text{tel} \}$

$\{ \text{sid}, \text{name} \} \rightarrow \{ \text{name}, \text{birthday}, \text{GPA}, \text{tel} \}$

Students

<u>sid</u>	name	birthday	GPA	tel
S001	陳會安	1967/9/3	3.7	02-22222222
S002	江小魚	1978/2/2	3.0	03-33333333
S003	張三丰	1982/3/3	3.2	04-44444444
S004	李四方	1981/4/4	2.9	05-55555555
S005	陳小安	1981/2/2	3.3	02-22222222
S006	江小魚	1976/9/9	3.5	06-66666666

- 如果一個屬性或一組屬性可以決定關聯表的所有屬性，這個屬性就是候選鍵。
 - PS. 應該是超鍵(Superkeys)
 - 例如：
- $\text{sid} \rightarrow \{ \text{name}, \text{birthday}, \text{GPA}, \text{tel} \}$

Students

<u>sid</u>	name	birthday	GPA	tel
S001	陳會安	1967/9/3	3.7	02-22222222
S002	江小魚	1978/2/2	3.0	03-33333333
S003	張三丰	1982/3/3	3.2	04-44444444
S004	李四方	1981/4/4	2.9	05-55555555
S005	陳小安	1981/2/2	3.3	02-22222222
S006	江小魚	1976/9/9	3.5	06-66666666

7-2-1 功能相依的定義-範例

- 例子: $R(A, B)$
- (X) $A \rightarrow B$
- (O) $B \rightarrow A$

A	B
1	4
1	5
3	7

7-2-2 功能相依的種類

22

- 功能相依的種類主要可以分為**三種**：
 1. 完全相依 (Full Dependency)
 2. 部分相依 (Partial Dependency)
 3. 遞移相依 (Transitive Dependency)

7-2-2 功能相依的種類-範例

23

- 關聯表範例是**Student_Course_Instructors**關聯表，如下圖所示：

Student_Course_Instructors

sid	name	c_no	title	eid	instructor	office
S001	陳會安	CS101	計算機概論	E001	陳慶新	CS-102
S001	陳會安	CS203	程式語言	E003	李鴻章	M-100
S001	陳會安	CS222	資料庫管理系統	E002	楊金欖	CIS-101
S001	陳會安	CS213	物件導向程式設計	E003	李鴻章	M-100
S002	江小魚	CS222	資料庫管理系統	E002	楊金欖	CIS-101
S002	江小魚	CS203	程式語言	E003	李鴻章	M-100
S003	張三丰	CS121	離散數學	E002	楊金欖	CIS-101
S003	張三丰	CS213	物件導向程式設計	E001	陳慶新	CS-102
S004	李四方	CS222	資料庫管理系統	E002	楊金欖	CIS-101

7-2-2 功能相依的種類-完全相依

24

- 完全相依是指如果屬性A（可能是複合屬性）和B是關聯表屬性，B功能相依於A，且B沒有功能相依於任何A的子集，則稱B完全功能相依於A。
- 例如：Student_Course_Instructors關聯表的完全相依，如下所示：

$\text{sid} \rightarrow \text{name}$

$\text{c_no} \rightarrow \text{title}$

$\text{eid} \rightarrow \text{instructor}$

$\text{eid} \rightarrow \text{office}$

$\{ \text{sid}, \text{c_no}, \text{eid} \} \rightarrow \{ \text{name}, \text{title}, \text{instructor}, \text{office} \}$

7-2-2 功能相依的種類-部分相依

25

- 部分相依是指如果A（可能是複合屬性）和B是關聯表的屬性，B是功能相依於A，如果刪除A子集的部分屬性，仍然不會影響功能相依，則稱B部分功能相依於A。
- 例如：`Student_Course_Instructors`關聯表的部分相依，如下所示：

`{ sid, c_no }→name`

`{ sid, c_no }→title`

`{ eid, instructor }→office`

7-2-2 功能相依的種類-遞移相依

26

- 如果A、B和C是關聯表的屬性， $A \rightarrow B$ 且 $B \rightarrow C$ ，則稱C是遞移相依於A， $A \rightarrow C$ 。
- 例如：Student_Course_Instructors關聯表的遞移相依，如下所示：

$\{ sid, c_no, eid \} \rightarrow office$

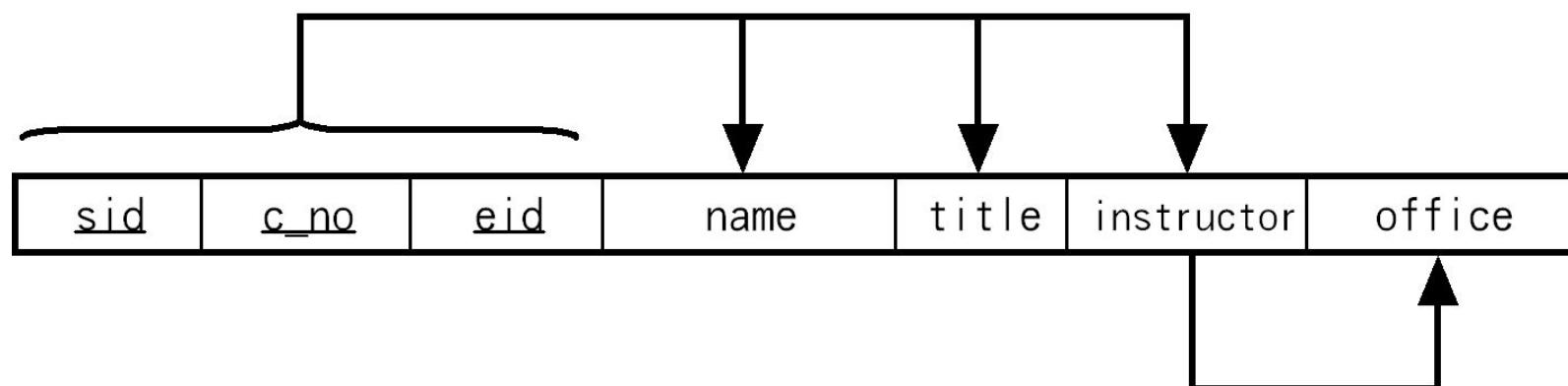
- 上述office屬性是遞移相依於 $\{ sid, c_no, eid \}$ ，因為：

$\{ sid, c_no, eid \} \rightarrow instructor$

$instructor \rightarrow office$

7-2-2 功能相依的種類-遞移相依

27



7-2-3 找出所有的功能相依

28

- **暴力演算法（Brute Force Algorithm）**：這是一種破解密碼常用的演算法，以作苦工方式嘗試所有可能的密碼組合；也就是說，在關聯表找出各屬性間的所有可能組合，然後一一檢查是否符合功能相依，如果關聯表的屬性很多，相對就需要花費更多的時間。
- **推論基礎演算法（Inference-based Algorithm）**：因為功能相依可以由其他功能相依來推論出，換句話說，推論基礎演算法是從關聯表已知的功能相依，使用推論規則推論出所有可能的功能相依。

7-2-4 阿姆斯壯推論規則-說明

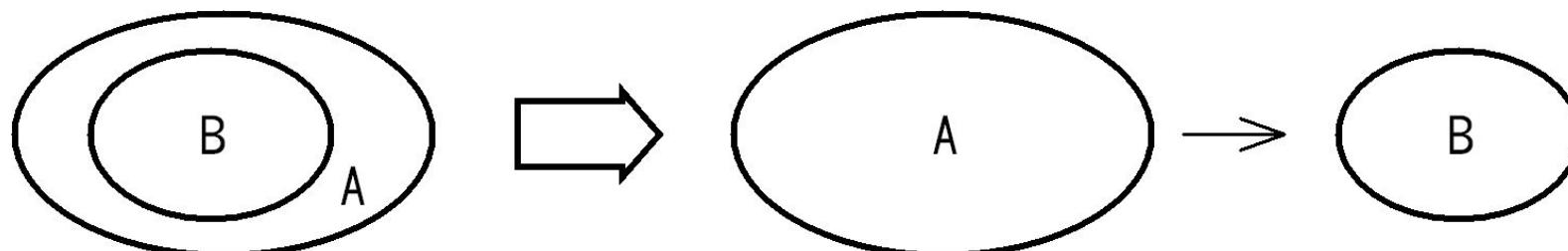
29

- **阿姆斯壯推論規則 (Armstrong's Inference Rule)**
可以從某些已知的功能相依推論出其他隱含的功能相依，這是一種**建全且完整**的規則，其意義如下：
 - **建全 (Sound)**：阿姆斯壯推論規則不會導出多餘的功能相依。
 - **完整 (Complete)**：阿姆斯壯推論規則可以導出所有的功能相依。

7-2-4 阿姆斯壯推論規則-反身規則 (Reflexitivity Rule)

30

- 如果屬性B是屬性A的子集合，則 $A \rightarrow B$ ，定義如下：
若： $\{ B_1, B_2, \dots, B_n \} \subseteq \{ A_1, A_2, \dots, A_n \}$
則： $\{ A_1, A_2, \dots, A_n \} \rightarrow \{ B_1, B_2, \dots, B_n \}$
- 例如： $\{ \text{sid}, \text{name} \} \rightarrow \{ \text{name} \}$
- 所有屬性都是本身的子集合，所以：
■ $\text{sid} \rightarrow \text{sid}$ 、 $\text{name} \rightarrow \text{name}$ 和 $\{ \text{sid}, \text{name} \} \rightarrow \{ \text{sid}, \text{name} \}$ 。



7-2-4 阿姆斯壯推論規則-擴充規則 (Augmentation Rule)

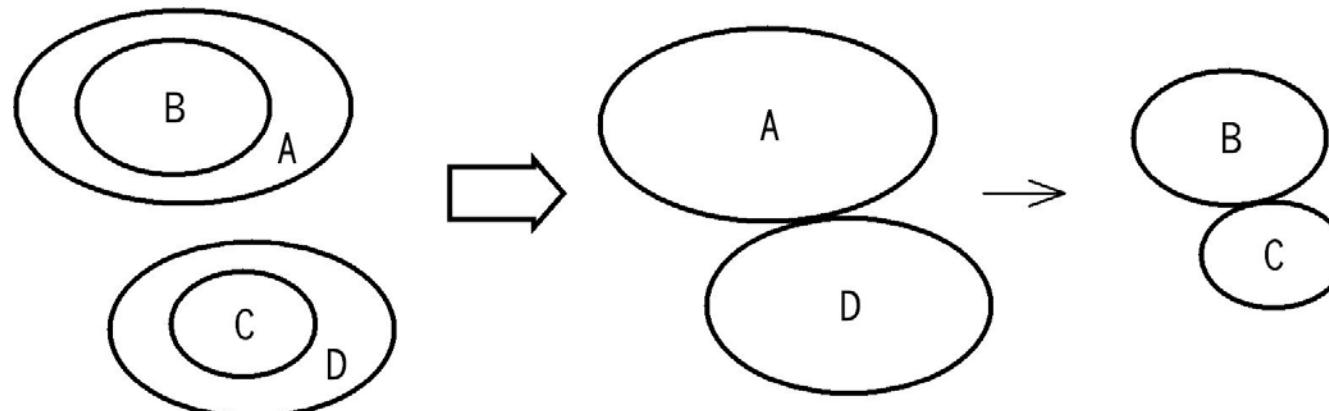
31

- 如果 $A \rightarrow B$ 且屬性 C 是屬性 D 的子集合，則 $AD \rightarrow BC$ ，其定義如下：

若： $\{A_1, A_2, \dots, A_n\} \rightarrow \{B_1, B_2, \dots, B_n\}$ 且 $\{C_1, \dots, C_n\} \subseteq \{D_1, \dots, D_n\}$

則： $\{A_1, A_2, \dots, A_n, D_1, \dots, D_n\} \rightarrow \{B_1, B_2, \dots, B_n, C_1, \dots, C_n\}$

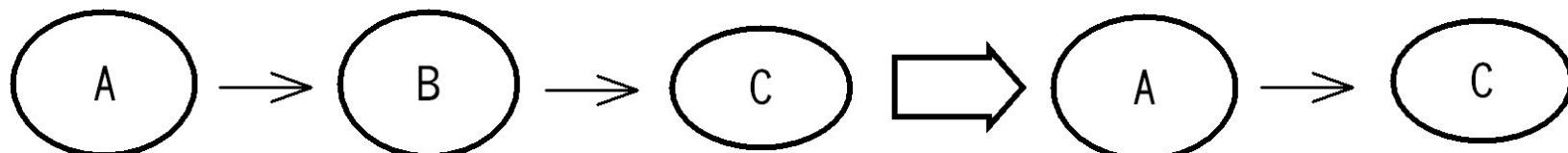
- 例如： $sid \rightarrow name$ 且 { GPA } 是 { birthday, GPA } 子集合，則可以得到：{ sid, birthday, GPA } \rightarrow { name, GPA }



7-2-4 阿姆斯壯推論規則-遞移規則 (Transitivity Rule)

32

- 如果 $A \rightarrow B$ 且 $B \rightarrow C$ ，則 $A \rightarrow C$ ，其定義如下：
若： $\{A_1, \dots, A_n\} \rightarrow \{B_1, \dots, B_n\}$ 且 $\{B_1, \dots, B_n\} \rightarrow \{C_1, \dots, C_n\}$
則： $\{A_1, \dots, A_n\} \rightarrow \{C_1, \dots, C_n\}$
- 例如： $sid \rightarrow instructor$ 且 $instructor \rightarrow office$ ，則可以得到：
 $sid \rightarrow office$



7-2-4 阿姆斯壯推論規則-分解規則 (Decomposition Rule)

33

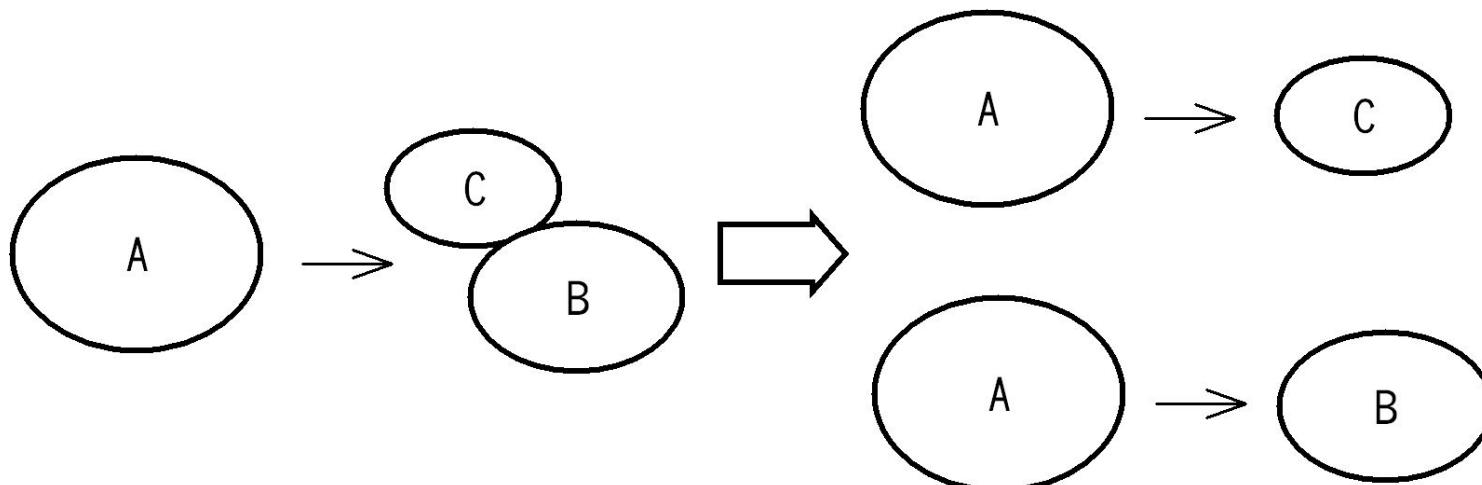
- 如果 $A \rightarrow BC$ ，則 $A \rightarrow B$ 且 $A \rightarrow C$ ，其定義如下：

若： $\{A_1, \dots, A_n\} \rightarrow \{B_1, \dots, B_n, C_1, \dots, C_n\}$

則： $\{A_1, \dots, A_n\} \rightarrow \{B_1, \dots, B_n\}$ 且 $\{A_1, \dots, A_n\} \rightarrow \{C_1, \dots, C_n\}$

- 例如： $sid \rightarrow \{name, birthday\}$ ，則可以得到：

$sid \rightarrow name$ 且 $sid \rightarrow birthday$



7-2-4 阿姆斯壯推論規則-聯集規則 (Union Rule)

34

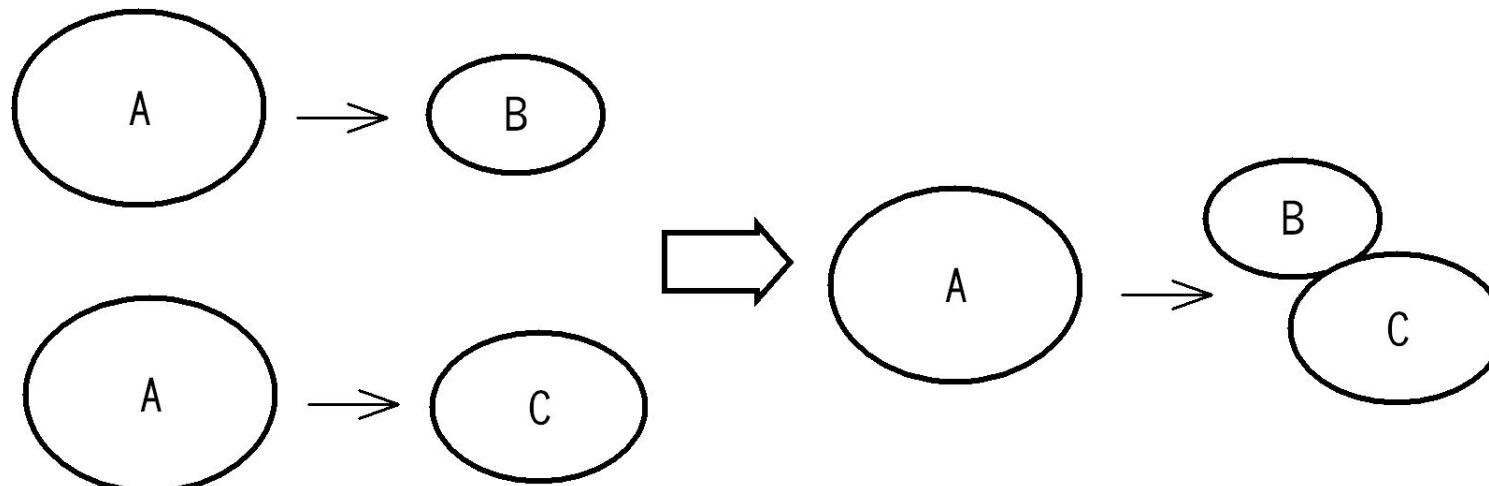
- 如果 $A \rightarrow B$ 且 $A \rightarrow C$ ，則 $A \rightarrow BC$ ，其定義如下：

若： $\{A_1, \dots, A_n\} \rightarrow \{B_1, \dots, B_n\}$ 且 $\{A_1, \dots, A_n\} \rightarrow \{C_1, \dots, C_n\}$

則： $\{A_1, \dots, A_n\} \rightarrow \{B_1, \dots, B_n, C_1, \dots, C_n\}$

- 例如： $sid \rightarrow name$ 且 $sid \rightarrow birthday$ ，則可得：

$sid \rightarrow \{name, birthday\}$



7-2-5 功能相依的推論範例-範例

35

- 現在只需使用推論規則，就可以從關聯表現有的功能相依推論出其他的功能相依。例如：關聯表已經擁有3條功能相依(FD)，如下所示：

1: $AB \rightarrow D$

2: $BD \rightarrow E$

3: $E \rightarrow C$

- 請使用上述3條功能相依推論出 $AB \rightarrow C$ 。

7-2-5 功能相依的推論範例-過程

步驟	使用規則	結果
1	$FD_1: AB \rightarrow D$	$FD_4: AB \rightarrow D$
2	$AB \subseteq AB$ 使用反身規則	$FD_5: AB \rightarrow AB$
3	以 $FD_5: AB \rightarrow AB$ 使用分解規則	$FD_6: AB \rightarrow B$
4	以 $FD_4: AB \rightarrow D$ 和 $FD_6: AB \rightarrow B$ 使用聯集規則	$FD_7: AB \rightarrow BD$
5	以 $FD_7: AB \rightarrow BD$ 和 $FD_2: BD \rightarrow E$ 使用遞移規則	$FD_8: AB \rightarrow E$
6	以 $FD_8: AB \rightarrow E$ 和 $FD_3: E \rightarrow C$ 使用遞移規則	$AB \rightarrow C$

7-2-5 功能相依的推論範例-改寫

步驟	使用規則	結果
1	因為 $B \subseteq AB$ 使用反身規則	$FD_4: AB \rightarrow B$
2	以 $FD_1: AB \rightarrow D$ 和 $FD_4: AB \rightarrow B$ 使用聯集規則	$FD_5: AB \rightarrow BD$
3	以 $FD_5: AB \rightarrow BD$ 和 $FD_2: BD \rightarrow E$ 使用遞移規則	$FD_6: AB \rightarrow E$
4	以 $FD_6: AB \rightarrow E$ 和 $FD_3: E \rightarrow C$ 使用遞移規則	Goal: $AB \rightarrow C$

7-2-5 功能相依的推論範例-另解

$FD_1: AB \rightarrow D$ $FD_2: BD \rightarrow E$ $FD_3: E \rightarrow C$

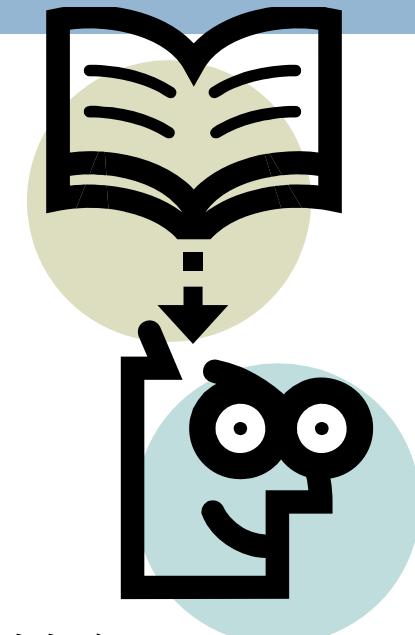
Goal: $AB \rightarrow C$

步驟	使用規則	結果
1	$FD_1: AB \rightarrow D$ 兩邊加上B (擴充規則)	$FD_4: AB \rightarrow BD$
2	$FD_4: AB \rightarrow BD$, $FD_2: BD \rightarrow E$ (遞移規則)	$FD_5: AB \rightarrow E$
3	以 $FD_5: AB \rightarrow E$ 和 $FD_3: E \rightarrow C$ (遞移規則)	Goal: $AB \rightarrow C$

7-3 第一階到第三階正規化型式

39

- 7-3-1 第一階正規化型式 – 1NF
- 7-3-2 第二階正規化型式 – 2NF
- 7-3-3 第三階正規化型式 – 3NF
- 7-3-4 Boyce-Codd正規化型式
 - BCNF
- 7-3-5 符合Boyce-Codd正規化型式的範例



7-3-1 第一階正規化型式 (1NF)-定義

40

定義7.2：關聯表R符合第一階正規化型式（First Normal Form；1NF）是指關聯表的每一個定義域（Domain）都是單元值（Atomic Value）且只能是單元值，也就是刪除多重值屬性型態（Multi-value Attribute Type）與複合屬性型態（Composite Attribute Type）的屬性。

- 簡單的說，上述定義是指關聯表沒有多重值和複合屬性。

7-3-1 第一階正規化型式 – 範例

41

- 例如：在**Students**關聯表儲存學生的選課資料，主鍵是學號**sid**，如下圖所示：

Students

<u>sid</u>	name	c_no	title	eid	instructor	office	room
S001	陳會安	{ CS101, CS203, CS222, CS213, }	{ 計算機概論, 程式語言, 資料庫管理系統, 物件導向程式設計, }	{ E001, E003, E002, E003, }	{ 陳慶新, 李鴻章, 楊金欖, 李鴻章 }	{ CS-102, M-100, CIS-101, M-100 }	{ 180-M, 221-S, 100-M, 500-K }
S002	江小魚	{ CS222, CS203 }	{ 資料庫管理系統, 程式語言 }	{ E002, E003 }	{ 楊金欖, 李鴻章 }	{ CIS-101, M-100 }	{ 100-M, 221-S }
S003	張三丰	{ CS121, CS213 }	{ 離散數學, 物件導向程式設計 }	{ E002, E001 }	{ 楊金欖, 陳慶新 }	{ CIS-101, CS-102 }	{ 221-S, 622-G }
S004	李四方	CS222	資料庫管理系統	E002	楊金欖	CIS-101	100-M

7-3-1 第一階正規化型式 – 正規化方法一

42

方法一：分割成不同的關聯表

- 關聯表違反1NF，一階正規化可以將這些多重值屬性連同主鍵分割成新關聯表，如下圖所示：

Students

sid	name
S001	陳會安
S002	江小魚
S003	張三丰
S004	李四方

Classes

sid	c_no	title	eid	instructor	office	room
S001	CS101	計算機概論	E001	陳慶新	CS-102	180-M
S001	CS203	程式語言	E003	李鴻章	M-100	221-S
S001	CS222	資料庫管理系統	E002	楊金欖	CIS-101	100-M
S001	CS213	物件導向程式設計	E003	李鴻章	M-100	500-K
S002	CS222	資料庫管理系統	E002	楊金欖	CIS-101	100-M
S003	CS203	程式語言	E003	李鴻章	M-100	221-S
S003	CS121	離散數學	E002	楊金欖	CIS-101	221-S
S003	CS213	物件導向程式設計	E001	陳慶新	CS-102	622-G
S004	CS222	資料庫管理系統	E002	楊金欖	CIS-101	100-M

7-3-1 第一階正規化型式 – 正規化方法二

43

方法二：分割成值組

- 1NF關聯表的每一個屬性只能儲存單元值，所以一階正規化可以將多重值屬性改成重複值組，屬性的每一個多重值新增一筆值組，如下圖所示：

Students

sid	name	c_no	title	eid	instructor	office	room
S001	陳會安	CS101	計算機概論	E001	陳慶新	CS-102	180-M
S001	陳會安	CS203	程式語言	E003	李鴻章	M-100	221-S
S001	陳會安	CS222	資料庫管理系統	E002	楊金欉	CIS-101	100-M
S001	陳會安	CS213	物件導向程式設計	E003	李鴻章	M-100	500-K
S002	江小魚	CS222	資料庫管理系統	E002	楊金欉	CIS-101	100-M
S002	江小魚	CS203	程式語言	E003	李鴻章	M-100	221-S
S003	張三丰	CS121	離散數學	E002	楊金欉	CIS-101	221-S
S003	張三丰	CS213	物件導向程式設計	E001	陳慶新	CS-102	622-G
S004	李四方	CS222	資料庫管理系統	E002	楊金欉	CIS-101	100-M

7-3-1 第一階正規化型式 – 正規化方法三

44

方法三：分割成不同屬性

- 一階正規化還可以將多重值屬性配合**虛值(Null)**新增為關聯表的多個屬性，不過其先決條件是**多重值的個數是有限的**。
- 例如：一位學生規定只能修兩門課程，**Students**關聯表可以使用2組屬性儲存選課資料的課程編號與名稱，雖然符合**1NF**，但是若學生選課數不只兩門，就會產生資料無法新增的異常情況。

1NF正規化方法三--多重值屬性配合null值 新增為關聯表的多個屬性

學號	姓名	手機1	手機2	手機3
S001	Tom	0921-000000	0922-000000	NULL
S002	John	0923-000000	NULL	NULL
S003	Mary	0924-000000	0925-000000	0926-000000

7-3-2 第二階正規化型式 (2NF)-定義

46

定義7.3：關聯表R符合第二階正規化型式 (Second Normal Form ; 2NF) 是指關聯表符合 1NF，而且所有非主鍵 (Primary Key) 的屬性都完全相依 (Fully Dependency) 於主鍵，也就是刪除所有部份相依的屬性。

- 簡單的說，上述定義是指關聯表中，不是主鍵的屬性需要完全相依於主鍵；反過來說，就是刪除關聯表所有部分相依 (Partial Dependency) 的屬性。

7-3-2 第二階正規化型式 -範例

47

- Students 關聯表執行一階正規化後，已經分割成 Students 和 Classes 關聯表，
- Classes 關聯表繼續第二階正規化，如下圖：

Classes

sid	c_no	title	eid	instructor	office	room
S001	CS101	計算機概論	E001	陳慶新	CS-102	180-M
S001	CS203	程式語言	E003	李鴻章	M-100	221-S
S001	CS222	資料庫管理系統	E002	楊金欉	CIS-101	100-M
S001	CS213	物件導向程式設計	E003	李鴻章	M-100	500-K
S002	CS222	資料庫管理系統	E002	楊金欉	CIS-101	100-M
S003	CS203	程式語言	E003	李鴻章	M-100	221-S
S003	CS121	離散數學	E002	楊金欉	CIS-101	221-S
S003	CS213	物件導向程式設計	E001	陳慶新	CS-102	622-G
S004	CS222	資料庫管理系統	E002	楊金欉	CIS-101	100-M

7-3-2 第二階正規化型式 – 功能相依

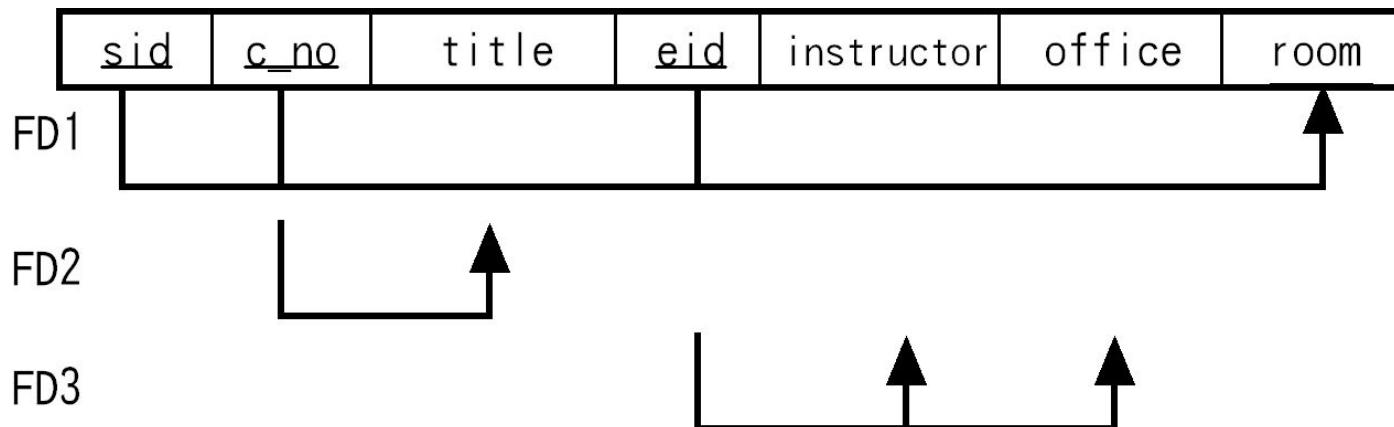
48

- Classes 關聯表的主鍵是 $(\text{sid}, \text{c_no}, \text{eid})$ ，關聯表已知的功能相依，如下所示：

FD1 : { sid, c_no, eid } \rightarrow room

FD2 : c_no \rightarrow title

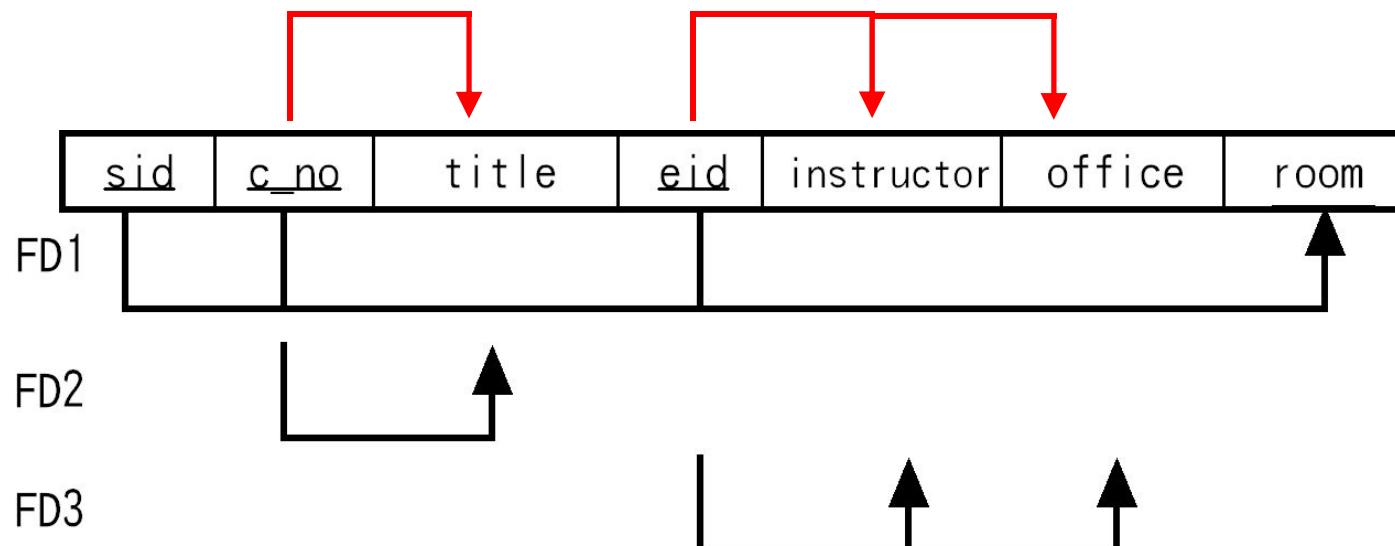
FD3 : eid \rightarrow { instructor, office }



7-3-2 第二階正規化型式 – 功能相依(續)

49

- 所以 **title** 是部分相依於主鍵 { sid, c_no, eid } ，因為 FD2 :
 $c_no \rightarrow title$
- 同理， **instructor, office** 也是部分相依於主鍵，因為
FD3 : $eid \rightarrow \{ instructor, office \}$



7-3-2 第二階正規化型式 – 正規化方式

50

- 部分相依表示有其他資料的子集合，應獨立成
新關聯表
 1. 將功能相依的**兩邊**的屬性獨立成新關聯表
 2. 功能相依的**左邊(決定屬性)**是新關聯表的主鍵
 3. 原關聯表必須有連外鍵(**FK**)指向新關聯表的主鍵

7-3-2 第二階正規化型式 – 正規化結果

51

Classes

<u>sid</u>	<u>c_no</u>	<u>eid</u>	room
S001	CS101	E001	180-M
S001	CS203	E003	221-S
S001	CS222	E002	100-M
S001	CS213	E003	500-K
S002	CS222	E002	100-M
S003	CS203	E003	221-S
S003	CS121	E002	221-S
S003	CS213	E001	622-G
S004	CS222	E002	100-M

Courses

<u>c_no</u>	title
CS101	計算機概論
CS203	程式語言
CS222	資料庫管理系統
CS213	物件導向程式設計
CS121	離散數學

Instructors

<u>eid</u>	instructor	office
E001	陳慶新	CS-102
E002	楊金欉	CIS-101
E003	李鴻章	M-100

7-3-3 第三階正規化型式 (3NF)-定義

52

定義7.4：關聯表R符合第三階正規化型式 (Third Normal Form ; 3NF) 是指關聯表符合2NF，而且所有非主鍵 (Primary Key) 的屬性都只能功能相依 (Functional Dependency) 於主鍵；沒有功能相依於其他非主鍵的屬性，即刪除遞移相依的屬性。

- 簡單的說，上述定義是指關聯表中不屬於主鍵的屬性都只能功能相依於主鍵，而不能同時功能相依於其他非主鍵的屬性，也就是刪除關聯表所有遞移相依 (Transitive Dependency) 的屬性。

7-3-3 第三階正規化型式 – 範例

53

- 例如：執行Instructors關聯表的第三階正規化，如下圖所示：

Instructors

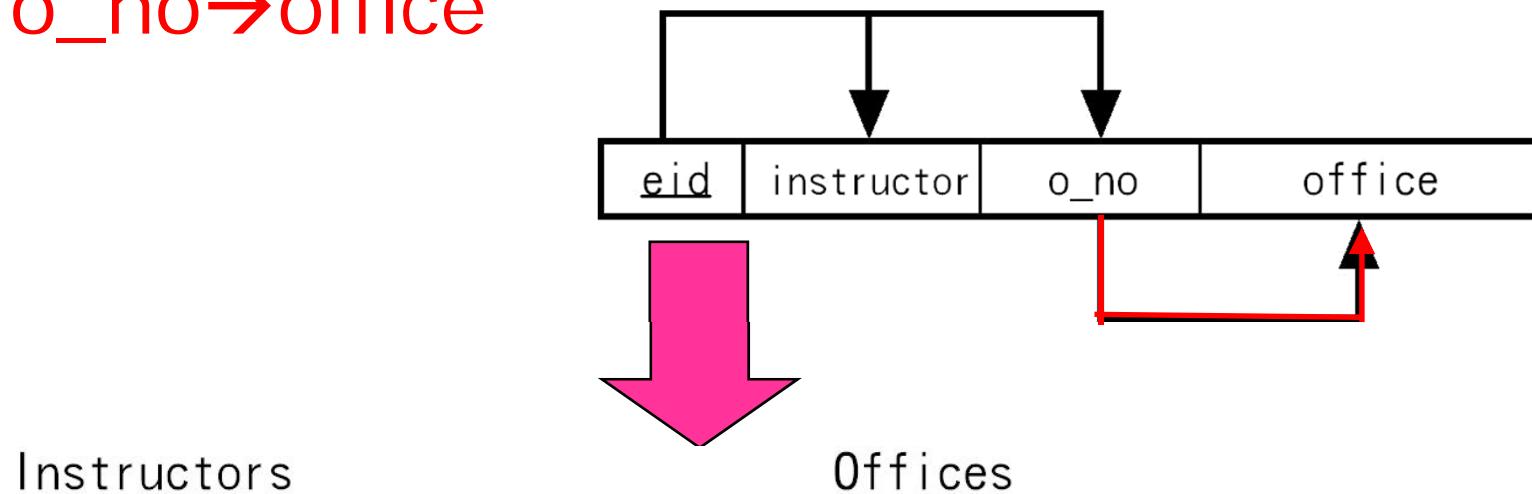


<u>eid</u>	instructor	o_no	office
E001	陳慶新	CS-102	網路研究室
E002	楊金欉	CIS-101	資料庫中心
E003	李鴻章	M-100	數學系電腦中心
E004	王陽明	M-100	數學系電腦中心

7-3-3 第三階正規化型式 – 正規化結果

54

- $\text{eid} \rightarrow \text{office}$ 是遞移相依，故須排除
 $\text{o_no} \rightarrow \text{office}$



Instructors

<u>eid</u>	instructor	<u>o_no</u>
E001	陳慶新	CS-102
E002	楊金欉	CIS-101
E003	李鴻章	M-100
E004	王陽明	M-100

Offices

<u>o_no</u>	office
CS-102	網路研究室
CIS-101	資料庫中心
M-100	數學系電腦中心

7-3-4 Boyce-Codd正規化型式 (BCNF) -條件

55

- Boyce-Codd正規化(BCNF)是在處理關聯表擁有多個候選鍵的特殊情況。Boyce-Codd正規化處理的關聯表至少擁有二個或更多個候選鍵，而且這兩個候選鍵是：
 1. 複合候選鍵
 2. 在複合候選鍵之間擁有重疊屬性，也就是說至少擁有一個相同屬性
- 如果關聯表沒有上述情況，3NF就等於BCNF。

7-3-4 Boyce-Codd正規化型式 – 定義

56

定義7.5：關聯表R符合BCNF正規化型式 (Boyce-Codd Normal Form ; BCNF) 是指關聯表所有主要的功能相依 $A \rightarrow B$, A一定且只可以是候選鍵 (Candidate Keys) , 也就是刪除只功能相依候選鍵部分屬性的功能相依。

- 簡單的說，BCNF是指關聯表中，主要功能相依 $A \rightarrow B$ 的左邊屬性A（稱為「決定屬性」，Determinant）一定是候選鍵或主鍵。
- 如下所示：候選鍵或主鍵 \rightarrow 其他屬性

7-3-4 Boyce-Codd 正規化型式 - 範例

57

- 例如：學生身份證字號（ SSN ）與成績（ grade ）的 Students 關聯表，如下圖所示：

Students

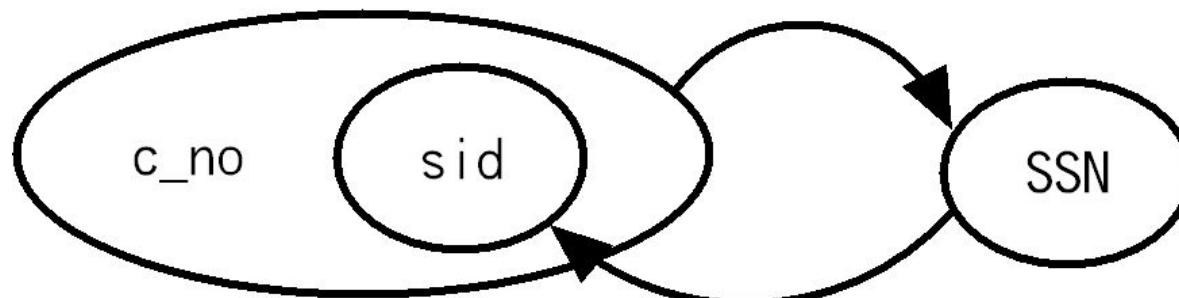
<u>sid</u>	SSN	<u>c_no</u>	grade
S001	H12345678	CS101	A
S001	H12345678	CS203	B
S002	J45678377	CS203	B
S003	I12345674	CS213	A-

- Students 關聯表擁有兩個候選鍵，如下所示：
(sid, c_no)
(SSN, c_no)

7-3-4 Boyce-Codd正規化型式 – 范例 (續)

58

- 兩個候選鍵擁有重疊屬性 c_no ，在候選鍵之間擁有所能相依 $SSN \rightarrow sid$ ，因為身份證字號可以決定學生編號，如下所示：
 $\{ sid, c_no \} \rightarrow SSN$ $SSN \rightarrow sid$
- 上述 SSN 屬性可以決定 sid 屬性，不過 SSN 只是候選鍵的一部分，且不是超鍵（ Superkey ），所以違反BCNF。



7-3-4 Boyce-Codd正規化型式 – 範例(續)

59

- Students符合2NF及3NF, 但是...
- 主鍵的一部分功能相依於非主鍵屬性



- R (sid, cno, SSN, grade)

7-3-4 Boyce-Codd正規化型式 – 正規化結果

60

- 例如：執行前述的關聯表Students關聯表的BCNF正規化，如下圖所示：

SSN

<u>sid</u>	SSN
S001	H12345678
S002	J45678377
S003	I12345674

Transcript

<u>sid</u>	<u>c_no</u>	grade
S001	CS101	A
S001	CS203	B
S002	CS203	B
S003	CS213	A-

7-3-4 Boyce-Codd正規化型式 – 正規化結果(續)

61

- 因為**Students**關聯表的**sid**和**SSN**兩個屬性間相互擁
有功能相依，如下所示：

$\text{sid} \rightarrow \text{SSN}$

$\text{SSN} \rightarrow \text{sid}$

- 所以執行BCNF正規化分割**Students**關聯表時，也
可以使用**SSN**屬性進行分割，如下圖所示：

SSN

<u>sid</u>	SSN
S001	H12345678
S002	J45678377
S003	I12345674

Transcript

<u>SSN</u>	<u>c_no</u>	grade
H12345678	CS101	A
H12345678	CS203	B
J45678377	CS203	B
I12345674	CS213	A-

7-3-5 符合Boyce-Codd正規化型式的範例

62

- Boyce-Codd正規化型式的條件是關聯表至少擁有一個或多個複合候選鍵。
- 不過關聯表擁有兩個複合候選鍵，並不表示一定需要進行BCNF正規化，兩個複合候選鍵的關聯表仍可以符合BNCF。

7-3-5 符合Boyce-Codd正規化型式的範例 (範例1)

63

- 例如：學生期末考座位的**Exams**關聯表，因為期末考分成多天進行考試，所以各科的考試座位可能不同，如下圖所示：

Exams

<u>sid</u>	<u>c_no</u>	seat_no
S001	CS101	C101-02
S001	CS203	C121-03
S002	CS203	C121-11
S003	CS213	C121-03

7-3-5 符合Boyce-Codd正規化型式的範例 (範例2)

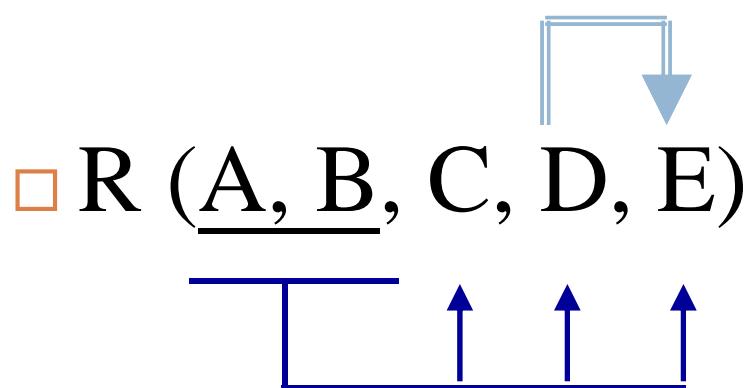
64

- 在Exams關聯表一共有兩個候選鍵，如下所示：
 $(\text{sid}, \text{c_no})$
 $(\text{seat_no}, \text{c_no})$
- 兩個候選鍵擁有重疊屬性 c_no ，不過在候選鍵之間並沒有功能相依 $\text{seat_no} \rightarrow \text{sid}$ 。
- 在Exams關聯表擁有的功能相依清單，如下所示：
 $\{ \text{sid}, \text{c_no} \} \rightarrow \text{seat_no}$
 $\{ \text{seat_no}, \text{c_no} \} \rightarrow \text{sid}$
- 上述2個功能相依的決定屬性都是候選鍵，符合BCNF定義。所以，Exams關聯表雖然擁有2組候選鍵，但是它不但符合3NF，而且符合BCNF。

複習-2NF

65

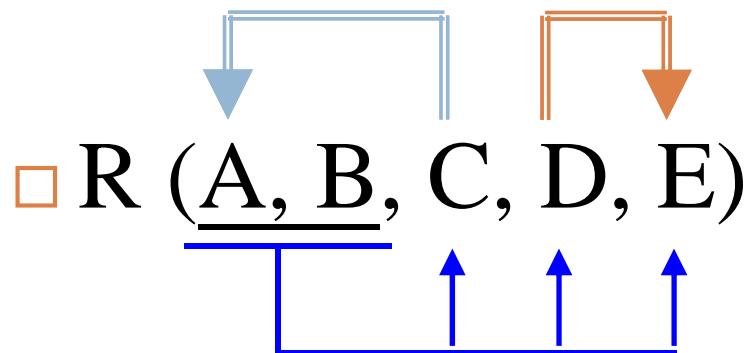
- 去除部分相依於PK
- $R(\underline{A, B}, C, D, E)$ 所有非鍵值屬性(例: C, D, E)必須功能相依於主鍵的全部($\{A, B\}$), 但不限制非鍵值屬性(E) 相依於其他非鍵值屬性(D)



複習-3NF

66

- 去除遞移相依
- 滿足2NF且所有非鍵值屬性(例:C, D, E)必須只功能相依於主鍵({A,B}), 但不限制非鍵值屬性(C)也能功能決定一些在候選鍵中的屬性(B)



不討論 trivial (瑣碎的, 普通的) FD:

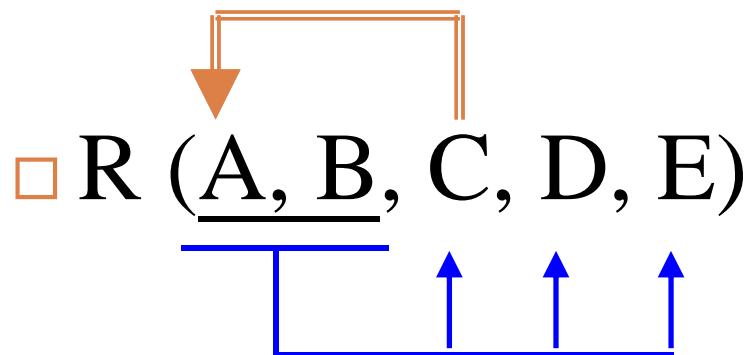
$S_1 \rightarrow S_2$ 若 $S_2 \subseteq S_1$,

例： $C \rightarrow C$

複習-BCNF(1)

67

- 滿足3NF, 且FD中所有決定屬性皆為候選鍵(超鍵),
但trivial FD不在討論範圍內!



複習-BCNF(2)

68

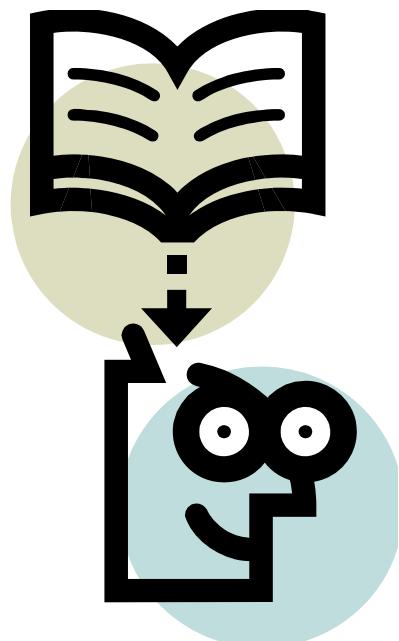
- 證明：若R有複合主鍵，且主鍵的一部分功能相依於非主鍵欄位(R違反BCNF)，則R：
 - 擁有二個或更多個候選鍵；
 - 複合候選鍵之間擁有重疊屬性
- 例：違反BCNF: R(A, B, C, D, E) 且 $C \rightarrow B$
 - $AC \rightarrow AB$ (擴充性)，
 - AC 為另一個候選鍵!(遞移性).
 - A為重疊屬性

練習

7-4 多重值相依與第四階正規化型式

69

- 7-4-1 多重值相依 (Multi-valued Dependency)
- 7-4-2 第四階正規化型式 – 4NF



7-4 多重值相依與第四階正規化型式

70

- 第四階正規化型式的基礎是「多重值相依」（**Multi-valued Dependency, MVD**），這是1977年R. Fagin所提出，屬於功能相依的廣義定義，即：功能相依是多重值相依的一種特例。
- 多重值相依的先決條件是關聯表擁有**3個***以上的屬性。
- *不盡然

7-4-1 多重值相依(範例)

71

- 例如：一門課可能由多位講師授課，每一門課可以使用多本教課書，儲存這些資料的 **Course_Instructor_Textbook** 關聯表，簡稱為 CIT，如下圖所示：

CIT

<u>course</u>	<u>instructor</u>	<u>textbook</u>
物件導向程式設計	陳慶新	Java 2程式設計範例教本
物件導向程式設計	陳慶新	物件導向程式設計
物件導向程式設計	李鴻章	Java 2程式設計範例教本
物件導向程式設計	李鴻章	物件導向程式設計
程式語言	李鴻章	C/C++程式設計範例教本
程式語言	李鴻章	Java 2程式設計範例教本
程式語言	李鴻章	VB.NET程式設計範例教本

7-4-1 多重值相依(說明)

72

- CIT關聯表的屬性仍然有類似功能相依的相依情況。例如：對於每一門課，存在課程相關教課書的子集合；同樣的，對於每一門課，存在此課程相關講師的子集合，如下所示：
 - 教課書textbook相依於課程course（在此的相依不是指功能相依）。
 - 講師instructor相依於課程course（在此的相依不是指功能相依）。
- 上述相依關係就是「多重值相依」（Multi-valued Dependency），以雙箭頭表示寫成：

course →→ textbook

course →→ instructor

7-4-1 多重值相依(定義)

73

定義7.6：多重值相依 (**Multi-valued Dependency**) 以關聯表 R 為例，若關聯表 R 是由屬性 A 、 B 與 C 組成 R(A, B, C) ， A 、 B 與 C 可以是複合屬性，我們可以說屬性 B 多重值相依 (**Multi-valued Dependent**) 於 A ，或 A 多重決定 (**Multi-determines**) 於 B ，寫成：

$R.A \rightarrow\!\!\! \rightarrow R.B$

或簡寫成：

$A \rightarrow\!\!\! \rightarrow B$

若關聯表 R 擁有 $A \rightarrow\!\!\! \rightarrow B$ 多重值相依，則表示存在 t_1 和 t_2 值組且 $t_1(A) = t_2(A)$ ，且另外存在二個值組 t_3 和 t_4 滿足下列條件：

$t_1(A) = t_2(A) = t_3(A) = t_4(A)$

$t_3(B) = t_1(B)$

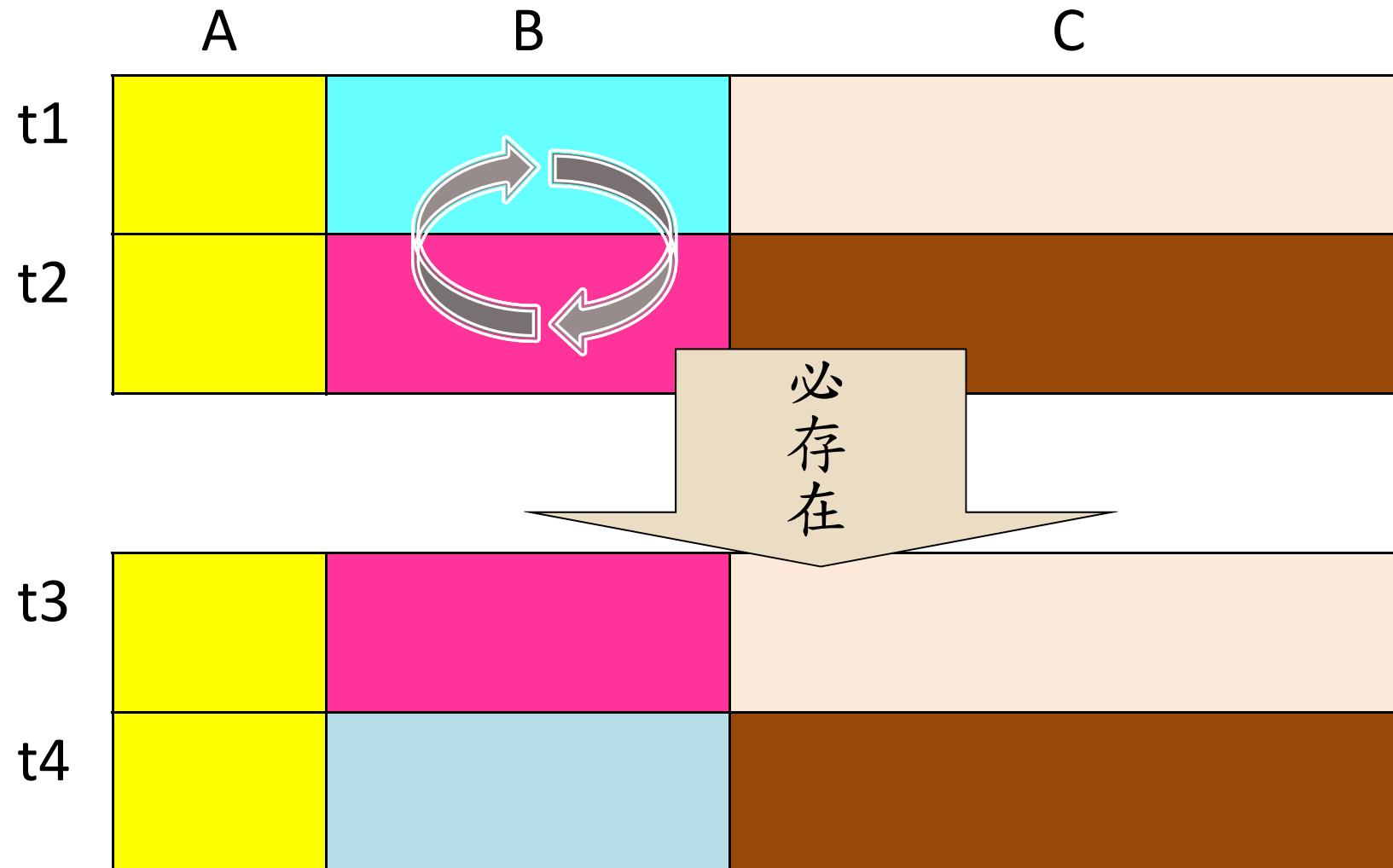
$t_3(C) = t_2(C)$

$t_4(B) = t_2(B)$

$t_4(C) = t_1(C)$

MVD: A →→ B

74



7-4-1 多重值相依(定義說明)

75

- 以CIT關聯表為例，因為關聯表CIT擁有 $\text{course} \rightarrow \rightarrow \text{textbook}$ ($A \rightarrow \rightarrow B$) 多重值相依，所以在關聯表可以找到交換 instructor 屬性的2對值組t1、t2和t3、t4，如下圖所示：

CIT		
	A	C
	<u>course</u>	<u>instructor</u>
t1	物件導向程式設計	陳慶新
t2	物件導向程式設計	李鴻章
t3	物件導向程式設計	李鴻章
t4	物件導向程式設計	陳慶新

7-4-1 多重值相依(異常狀況)

76

- 在CIT關聯表擁有很多重複資料，這些重複資料會導致新增或更新異常。例如：新增講師陳慶新教授程式語言這門課，雖然只有一門課，但是需要同時新增3筆值組，每一個值組是一本教課書，否則就會產生資料不一致的問題。

CIT

course	instructor	textbook
物件導向程式設計	陳慶新	Java 2程式設計範例教本
物件導向程式設計	陳慶新	物件導向程式設計
物件導向程式設計	李鴻章	Java 2程式設計範例教本
物件導向程式設計	李鴻章	物件導向程式設計
程式語言	李鴻章	C/C++程式設計範例教本
程式語言	李鴻章	Java 2程式設計範例教本
程式語言	李鴻章	VB.NET程式設計範例教本
程式語言	陳慶新	C/C++程式設計範例教本
程式語言	陳慶新	Java 2程式設計範例教本
程式語言	陳慶新	VB.NET程式設計範例教本

7-4-1 多重值相依(推論規則1)*

77

- 多重值相依的推論規則是1977年C. Beeri、R. Fagin 和J. H. Howard擴充自阿姆斯壯推論規則，其中前3個推論規則是源自功能相依，如下所示：
 - **FD反身規則（FD Reflexitivity）**：如果屬性B是屬性A的子集合，則 $A \rightarrow B$ 。
 - **FD擴充規則（FD Augmentation）**：如果 $A \rightarrow B$ 且屬性C是屬性D的子集合，則 $AD \rightarrow BC$ 。
 - **FD遞移規則（FD Transitivity）**：如果 $A \rightarrow B$ 且 $B \rightarrow C$ ，則 $A \rightarrow C$ 。
 - **MVD擴充規則（MVD Augmentation）**：如果 $A \rightarrow\rightarrow B$ 且屬性C是屬性D的子集合，則 $AD \rightarrow\rightarrow BC$ 。
 - **MVD遞移規則（MVD Transitivity）**：如果 $A \rightarrow\rightarrow B$ 且 $B \rightarrow\rightarrow C$ ，則 $A \rightarrow\rightarrow C - B$ 。

7-4-1 多重值相依(推論規則2)*

78

- **MVD互補規則 (MVD Complementation)**：如果 $A \rightarrow\rightarrow B$ ，則 $A \rightarrow\rightarrow (R - A - B)$ 。
- **MVD聯集規則 (MVD Union)**：如果 $A \rightarrow\rightarrow B$ 且 $A \rightarrow\rightarrow C$ ，則 $A \rightarrow\rightarrow BC$ 。
- **MVD分解規則 (MVD Decomposition)**：如果 $A \rightarrow\rightarrow BC$ ，則 $A \rightarrow\rightarrow B C$ ， $A \rightarrow\rightarrow B - C$ 且 $A \rightarrow\rightarrow C - B$ 。
- **複製規則 (Replication或稱為升級Promotion)**：如果 $A \rightarrow B$ ，則 $A \rightarrow\rightarrow B$ 。
- **合併規則 (Coalescence)**：如果 $A \rightarrow\rightarrow B$ 且 $C \rightarrow D$ ， D 是 B 的子集合， C 與 B 的交集是空集合，則 $A \rightarrow D$ 。
- **偽裝遞移規則 (Pseudotransitivity)**：如果 $A \rightarrow\rightarrow B$ 且 $CB \rightarrow\rightarrow D$ ，則 $CA \rightarrow\rightarrow D - CB$ 。
- **混合偽裝遞移規則 (Mixed Pseudotransitivity)**：如果 $A \rightarrow\rightarrow B$ 且 $AB \rightarrow\rightarrow C$ ，則 $A \rightarrow\rightarrow C - B$ 。

7-4-2 第四階正規化型式 – 4NF(定義)

79

- 定義7.7：關聯表R符合第四階正規化型式（Fourth Normal Form；4NF）是指關聯表符合BCNF，而且所有多重值相依（Multi-valued Dependency）都是功能相依，也就是刪除多重值相依。
- 簡單的說，上述定義是指關聯表沒有存在的多重值相依，只有功能相依。

7-4-2 第四階正規化型式 (4NF) - 正規化

80

- CIT關聯表擁有2個多重值相依，如下所示：

`course`→→`textbook`

`course`→→`instructor`

- 執行第四階正規化，只需將多重值相依A→→B兩邊的屬性獨立成關聯表

CI

<u>course</u>	<u>instructor</u>
物件導向程式設計	陳慶新
物件導向程式設計	李鴻章
程式語言	李鴻章

CIT	course	instructor	textbook
物件導向程式設計	陳慶新	Java 2程式設計範例教本	
物件導向程式設計	陳慶新	物件導向程式設計	
物件導向程式設計	李鴻章	Java 2程式設計範例教本	
物件導向程式設計	李鴻章	物件導向程式設計	
程式語言	李鴻章	C/C++程式設計範例教本	
程式語言	李鴻章	Java 2程式設計範例教本	
程式語言	李鴻章	VB.NET程式設計範例教本	
程式語言	陳慶新	C/C++程式設計範例教本	
程式語言	陳慶新	Java 2程式設計範例教本	
程式語言	陳慶新	VB.NET程式設計範例教本	

4NF補充範例-1

81

- R(貸款編號, 客戶, 居住地區, 居住城市)
- 客戶 →→ 居住地區, 居住城市

貸款編號	客戶	居住地區	居住城市
L-23	Bob	西屯區	台中
L-23	Bob	信義區	台北
L-93	Alice	北屯區	台中

4NF補充範例-2

82

- 客戶 →→ 居住地區, 居住城市

客戶	居住地區	居住城市	貸款編號
Bob	西屯區	台中	L-23
Bob	信義區	台北	L-23
Bob	西屯區	台中	L-25
Bob	信義區	台北	L-25

4NF補充範例-3

83

□ 客戶 →→ 貸款編號

客戶	貸款編號	居住地區	居住城市
Bob	L-23	西屯區	台中
Bob	L-23	信義區	台北
Bob	L-25	西屯區	台中
Bob	L-25	信義區	台北

4NF補充範例-4

84

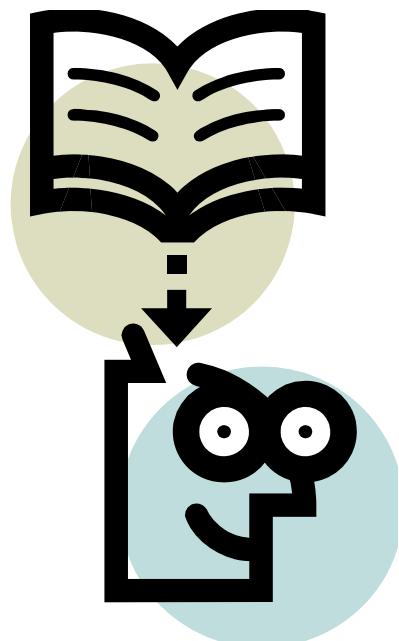
- R(貸款編號, 客戶, 居住地區, 居住城市)
- 客戶 →→ 居住地區, 居住城市
 - 或 : 客戶 →→ 貸款編號

- 分割成 :
- R1(客戶, 居住地區, 居住城市)
- R2(客戶, 貸款編號)

*7-5 合併相依與第五階正規化型式

85

- 7-5-1 合併相依 (Join Dependency)
- 7-5-2 第五階正規化型式 – 5NF



7-5 合併相依與第五階正規化型式

86

- 關聯式資料庫正規化分析的第一階到第三階正規化型式是使用功能相依找出關聯表的重複資料。
- 在第四階正規化型式的基礎是多重值相依，這是功能相依的擴充。
- 第五階正規化型式的基礎則是擴充多重值相依的「合併相依」（Join Dependency）。
- 簡單的說，合併相依是指當關聯表分割成3個或更多關聯表後，一定能夠透過多次合併運算恢復成原來的關聯表。

7-5-1 合併相依(範例)

87

- 例如：每個科系（**department**）開多門課，課程（**course**）可以給多位學生修，學生（**student**）可以修不同科系的課，這3個屬性循環關聯儲存在同一個關聯表DCS，如下圖所示：

DCS

<u>depar tment</u>	<u>course</u>	<u>student</u>
資訊系	101	陳會安
資訊系	202	江小魚
資管系	202	江小魚
資管系	101	張三丰
資管系	222	江小魚

7-5-1 合併相依(說明)

88

- 將DCS分割成2個關聯表，看看是否可以解決資料重複的問題，分割方式有多種組合，如下所示：
DC (department, course) 、 CS (course, student)
- 或：
DC (department, course) 、 DS (department, student)
- 以第一種分割方式為例，在分割成2個關聯表DC和CS後，可以看到減少一些資料重複，不過如果將這2個關聯表以自然合併進行結合，可以發現多出幾筆值組，和原始關聯表不同。

7-5-1 合併相依(自然合併)

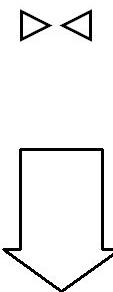
89

DC

<u>department</u>	<u>course</u>
資訊系	101
資訊系	202
資管系	202
資管系	101
資管系	222

CS

<u>course</u>	<u>student</u>
101	陳會安
202	江小魚
101	張三丰
222	江小魚



DCS

<u>department</u>	<u>course</u>	<u>student</u>
資訊系	101	陳會安
資訊系	101	張三丰
資訊系	202	江小魚
資管系	202	江小魚
資管系	101	陳會安
資管系	101	張三丰
資管系	222	江小魚

CH5-3-2合併
運算

(參考) 5-3-2 合併運算-說明

90

- 合併運算是在兩個關聯表中，使用具有相同定義域（Domain）的屬性為條件合併兩個關聯表，相當於先執行基本運算的卡笛生乘積運算，再配合選擇運算取出所需值組的關聯表。
- 合併運算滿足交換律和結合律，合併運算可以分為數種：
 - θ 合併
 - Equijoin合併
 - 自然合併
 - 外部合併

(參考)5-3-2 合併運算- 自然合併(語法)

91

- 自然合併是一種使用主鍵和外來鍵為條件的合併方式，其隱含的條件是主鍵等於外來鍵條件的 Equijoin 合併，所以自然合併屬於一種特殊格式的 Equijoin 合併，其基本語法和定義如下所示：

$$R \bowtie S = \pi_{r \cup s}(R \bowtie_{r=s} S)$$

- 上述語法並沒有指明條件，這是因為預設使用主鍵和外來鍵對應的條件，即所有共通屬性值（All Common Attributes）都需要相等。
- 自然合併的執行結果關聯表綱要的共通屬性部分在運算結果中只會顯示一次。

(參考) 5-3-2 合併運算- 自然合併(範例)

92

- 例如：Students和Departments關聯表的自然合併運算式，如下所示：

Result = Students  Departments

Result

sid	Students.name	d_no	Departments.name	room
S001	陳會安	CS	資訊系	CS-1
S002	江小魚	CIS	資管系	B-21

```
SELECT * FROM Students INNER JOIN Departments  
ON Students.d_no = Departments.d_no
```

7-5-1 合併相依(投影分割)

93

- 將關聯表使用投影運算分割成3個關聯表：DC (*department, course*) 、 CS (*course, student*) 、 SD (*student, department*) ，如下圖所示：

DC

<u>department</u>	<u>course</u>
資訊系	101
資訊系	202
資管系	202
資管系	101
資管系	222

CS

<u>course</u>	<u>student</u>
101	陳會安
202	江小魚
101	張三丰
222	江小魚

SD

<u>student</u>	<u>department</u>
陳會安	資訊系
江小魚	資訊系
江小魚	資管系
張三丰	資管系

7-5-1 合併相依(自然合併)

94

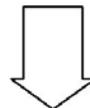
- 關聯表可以分割成3個或更多個關聯表，然後透過多次合併運算恢復成原始關聯表，就稱為「合併相依」（Join Dependency）。

DCS_Temp

department	course	student
資訊系	101	陳會安
資訊系	101	張三丰
資訊系	202	江小魚
資管系	202	江小魚
資管系	101	陳會安
資管系	101	張三丰
資管系	222	江小魚

SD

student	department
陳會安	資訊系
江小魚	資訊系
江小魚	資管系
張三丰	資管系



DCS

department	course	student
資訊系	101	陳會安
資訊系	202	江小魚
資管系	202	江小魚
資管系	101	張三丰
資管系	222	江小魚

7-5-1 合併相依(定義)

95

定義7.8：合併相依 (Join Dependency) 以關聯表 R 為例，若關聯表 R 等於自然合併運算 R_1, R_2, \dots, R_n ，
 R_i 是關聯表 R 屬性的子集合，我們可以說，
關聯表 R 滿足合併相依 (R_1, R_2, \dots, R_n) 。

- 關聯表 DCS 可以使用投影運算分割成 DC 、 CS 和 SD 三個關聯表，執行 $DC \bowtie CS \bowtie SD$ 自然合併運算就可以還原成原始 DCS 關聯表，所以關聯表 DCS 擁有合併相依。

7-5-1 合併相依(異常狀況1)

96

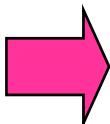
- 如果關聯表擁有合併相依，在新增和刪除時就會產生異常情況，例如：DCS關聯表擁有合併相依，如下圖所示：

DCS

<u>department</u>	<u>course</u>	<u>student</u>
資訊系	101	陳會安
資訊系	202	江小魚

在DCS關聯表新增一筆值組{資管系, 101, 江小魚}，如右圖所示：

DCS



<u>department</u>	<u>course</u>	<u>student</u>
資訊系	101	陳會安
資訊系	202	江小魚
資管系	101	江小魚

7-5-1 合併相依(異常狀況2)

97

- 將上述關聯表以投影運算分割成DC、CS和SD三個關聯表，再將它使用合併運算結合起來，可以發現多了一筆值組，如下圖所示：

DCS

department	course	student
資訊系	101	陳會安
資訊系	101	江小魚
資訊系	202	江小魚
資管系	101	江小魚

- 如果新增一個值組，例如資訊系101江小魚，在分割和合併後，就不會多出一筆值組。

7-5-2 第五階正規化型式 – 5NF(定義)

98

定義7.9：關聯表R符合第五階正規化型式（**Fifth Normal Form ; 5NF**），也稱為投影-合併正規化型式（**Project-Join Normal Form ; PJNF**），這是指關聯表符合4NF，且所有合併相依(R_1, R_2, \dots, R_n)的關聯表屬性子集 R_i ，都是關聯表R的超鍵（**Superkey**），也就是刪除合併相依。

- 上述定義是指關聯表沒有存在合併相依。以上一節的DCS關聯表為例，DCS關聯表擁有合併相依，所以不符合5NF。

7-5-2 第五階正規化型式 – 5NF(正規化)

99

- DCS關聯表擁有合併相依，可以使用投影運算進行再次的分割，並不符合5NF，所以需要執行第五階正規化，使用投影運算分割成3個關聯表，如下圖所示：

DC

<u>department</u>	<u>course</u>
資訊系	101
資訊系	202
資管系	202
資管系	101
資管系	222

CS

<u>course</u>	<u>student</u>
101	陳會安
202	江小魚
101	張三丰
222	江小魚

SD

<u>student</u>	<u>department</u>
陳會安	資訊系
江小魚	資訊系
江小魚	資管系
張三丰	資管系

練習

100

- 紿定關聯表R和已知的功能相依分別如下：
 - a) $R(\underline{A}, B, C, D, E)$, FD: $B \rightarrow C, D \rightarrow E$
 - b) $R(\underline{A}, B, C, D, E, F)$, FD: $EF \rightarrow B, D \rightarrow C$
 - c) $R(\underline{A}, B, C, D)$, FD: $A \rightarrow B, BC \rightarrow D, A \rightarrow C$
- 請分別分割(a)(b)(c)中的關聯表R，得到一組皆符合**BCNF**的關聯表綱要（需標明主鍵）。

