

第5章 關聯式資料庫的正規化

- ◆ 5-1 正規化的基礎
- ◆ 5-2 功能相依
- ◆ 5-3 第一階到第三階正規化型式



5-1 正規化的基礎-說明

◆ 正規化（Normalization）

- ◆ 一種標準處理程序
- ◆ 決定關聯表應該有哪些屬性
- ◆ 目的是建立良好結構關聯表(Well-structured Relation)

◆ 良好結構關聯表

- ◆ 一種沒有重複資料的關聯表
- ◆ 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

5-1 正規化的基礎-目的

◆ 去除資料重複

- ◆ 建立儘可能沒有重複資料的關聯表
- ◆ 重複資料缺點
 - ◆ 浪費資料庫的儲存空間
 - ◆ 產生資料維護上的問題

◆ 去除不良的資料相依

- ◆ 資料相依是指關聯表中的屬性之間有相依關係
- ◆ 如果關聯表有不良的資料相依，在新增、刪除、更新資料時，會造成異常現象

5-1-1 關聯表為什麼需要正規化-說明

- ◆ 關聯表有非自然結合的屬性
 - ◆ 稱為非正規化型式（Unnormalized Form）關聯表
 - ◆ 會造成資料重複的問題



5-1-1 關聯表為什麼需要正規化- 更新異常(Update Anomaly)

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



5-1-1 關聯表為什麼需要正規化- 刪除異常(Deletion Anomaly)

- ◆ 在Student_Course_Classes關聯表刪除值組
 - ◆ 可能刪除合法資料而無法再取得
 - ◆ 例如CS101和CS121課程都只有一筆值組



5-1-1 關聯表為什麼需要正規化- 新增異常(Insertion Anomaly)

- ◆ 在Student_Course_Classes關聯表新增值組
 - ◆ 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



5-1-2 正規化的型式-階段

- ◆ 正規化的處理過程
 - ◆ 分為五個階段的正規化型式
 - ◆ 每一階段的正規化型式，處理不同的屬性間資料相依的問題



5-1-2 正規化的型式-前三階

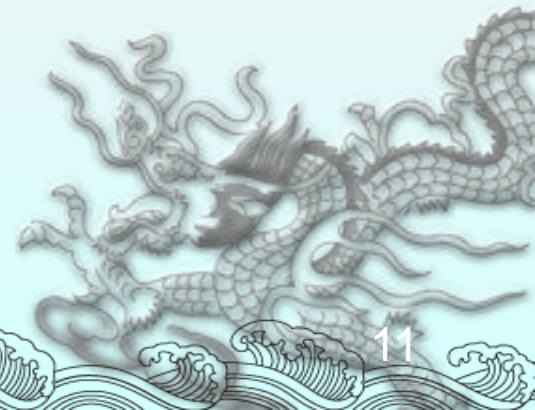
- ◆ 第一正規化型式(First Normal Form, **1NF**)
 - ◆ 關聯表刪除多重值屬性
 - ◆ 關聯表只有單元值屬性
- ◆ 第二正規化型式(Second Normal Form, **2NF**)
 - ◆ 關聯表滿足1NF且
 - ◆ 關聯表沒有部分相依 (Partial Dependency)
- ◆ 第三正規化型式(Third Normal Form, **3NF**)
 - ◆ 關聯表滿足2NF且
 - ◆ 關聯表沒有遞移相依 (Transitive Dependency)

5-1-2 正規化的型式-BCNF+四和五階

- ◆ **Boyce-Codd正規化型式**(Boyce-Codd Normal Form, **BCNF**)
 - ◆ 一種廣意的第三正規化型式
 - ◆ 如果關聯表有多個複合候選鍵，刪除**候選鍵**之間的功能相依
- ◆ 第四正規化型式(Fourth Normal Form, 4NF)
 - ◆ 關聯表滿足BCNF且沒有**多重值相依**
- ◆ 第五正規化型式(Fifth Normal Form, 5NF)
 - ◆ 關聯表滿足4NF且沒有**合併相依**

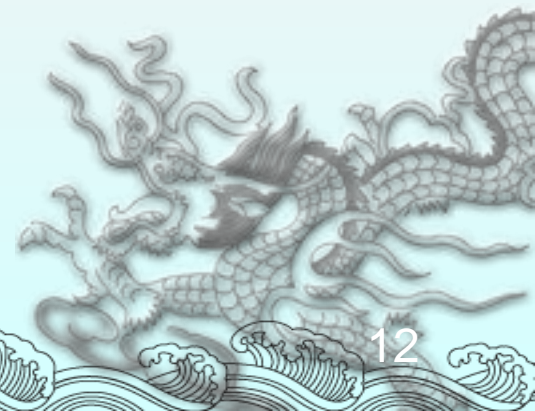
5-2 功能相依-說明

- ◆ 功能相依（Functional Dependency，簡稱FD）
 - ◆ 描述關聯表中屬性之間的相依關係
 - ◆ 關聯式資料庫正規化的基礎
 - ◆ 關聯表正規化的第一步，就是找出屬性之間的功能相依



5-2 功能相依-目的

- ◆ 找出屬性之間的功能相依，就可以找出
 - ◆ 重複資料：一些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(組)屬性就是候選鍵或主鍵



5-2-1 功能相依-定義

◆ 功能相依定義

- ◆ 若關聯表R有屬性A與B，A與B可以是複合屬性
- ◆ 如果B功能相依於A，或稱A功能決定B
- ◆ 則寫成： $R.A \rightarrow R.B$ ，或簡寫成： $A \rightarrow B$

5-2-1 功能相依-說明

- ◆ 關聯表中，若存在功能相依 $A \rightarrow B$
 - ◆ 表示知道屬性集A的值，就可以知道屬性集B的值
 - ◆ 屬性集A可能是 $\{A1\}$ 或 $\{A1, A2, \dots, An\}$
 - ◆ 屬性集B可能是 $\{B1\}$ 或 $\{B1, B2, \dots, Bn\}$
 - ◆ 所以下列式子有可能出現
 - $A1 \rightarrow B1$
 - $A1 \rightarrow \{B1, B2, \dots, Bn\}$
 - $\{A1, A2, \dots, An\} \rightarrow B1$
 - $\{A1, A2, \dots, An\} \rightarrow \{B1, B2, \dots, Bn\}$
 - ◆ 功能相依屬於一種單向的關聯
 - ◆ $A \rightarrow B$ 不表示 $B \rightarrow A$

5-2-1 功能相依-範例

◆ 在Students關聯表的一些功能相依範例

$sid \rightarrow name$ $sid \rightarrow birthday$ $sid \rightarrow tel$

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

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

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

5-2-2 功能相依-種類

- ◆ 功能相依的種類主要分為三種
 - ◆ 完全相依(Full Dependency)
 - ◆ 部分相依(Partial Dependency)
 - ◆ 遞移相依(Transitive Dependency)
- ◆ 範例：Student_Course_Instructors關聯表

5-2-2 功能相依-完全相依

◆ 完全相依

- ◆ A和B是關聯表屬性，A或B可能是複合屬性
- ◆ 如果B功能相依於A，而且
- ◆ B沒有功能相依於任何A的子集
- ◆ 則稱B完全相依於A
- ◆ Student_Course_Instructors關聯表的完全相依
 $\{ \text{sid}, \text{c_no}, \text{eid} \} \rightarrow \{ \text{name}, \text{title}, \text{instructor}, \text{office} \}$

5-2-2 功能相依的種類-部分相依

◆ 部分相依

- ◆ A(複合屬性, 例如a1+a2)和B是關聯表的屬性
- ◆ 如果B是功能相依於A ($A \rightarrow B$), 而且
- ◆ B又功能相依於A的部分屬性 ($a1 \rightarrow B$ 或 $a2 \rightarrow B$)
- ◆ 則稱B部分相依於A
- ◆ Student_Course_Instructors關聯表的部分相依
 - { sid, c_no, eid } \rightarrow name (因為 $sid \rightarrow name$)
 - { sid, c_no, eid } \rightarrow title (因為 $c_no \rightarrow title$)
 - { sid, c_no, eid } \rightarrow instructor (因為 $eid \rightarrow instructor$)

5-2-2 功能相依的種類-遞移相依

◆ 遞移相依

- ◆ A、B和C是關聯表的屬性
- ◆ 如果 $A \rightarrow B$ 且 $B \rightarrow C$ ，則稱C是遞移相依於A， $A \rightarrow C$
- ◆ Student_Course_Instructors關聯表的遞移相依
 { sid, c_no, eid } \rightarrow office 因為
 { sid, c_no, eid } \rightarrow instructor 而 instructor \rightarrow office



5-3-1 第一正規化型式 – 1NF(定義)

- ◆ 關聯表R符合第一正規化型式（1NF）
 - ◆ 每一個屬性值都是單元值
 - ◆ 須有一個PK屬性，使其他屬性功能相依於此PK



5-3-1 第一正規化型式 – 1NF(範例)

- ◆ 在Students資料表儲存學生的選課資料，主鍵是學號sid

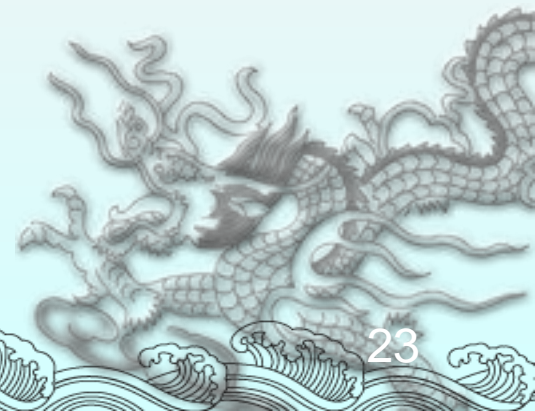


5-3-1 第一正規化型式 – 1NF(方法)

- ◆ 分割成不同的關聯表
 - ◆ 關聯表Students違反1NF
 - ◆ 將多重值屬性連同主鍵分割成新關聯表

5-3-2 第二正規化型式 – 2NF(定義)

- ◆ 關聯表R符合第二正規化型式（2NF）
 - ◆ 關聯表符合1NF
 - ◆ 所有非主鍵的屬性都只能功能相依於主鍵
 - ◆ 沒有部分相依情況發生



5-3-2 第二正規化型式 – 2NF(範例)

- ◆ Students關聯表執行第一正規化後，已經分割成Students和Classes關聯表
- ◆ Classes關聯表繼續第二正規化，因為有部分相依情況



5-3-2 第二正規化型式 – 2NF(功能相依)

- ◆ Classes關聯表的主鍵是(sid, c_no, eid)，已知的功能相依
FD1 : { sid, c_no, eid } → { title, instructor, office, room }
FD2 : c_no → title
FD3 : eid → { instructor, office }
FD4 : instructor → office

補充：1NF轉2NF三步驟

- ◆ 將PK的各個子集合欄位，依欄位數量由小到大排列出來，每個子集合自成一列，視為一個新PK。
- ◆ 由上而下檢查各個新PK，將功能相依於此新PK的欄位列在其後面，已列出過的欄位不可重複出現。
- ◆ 刪除只有新PK而無其他欄位的列，將剩下來的每一列視為一個關聯表綱要，取適當名稱，並標示其PK。

補充：1NF轉2NF例子

◆ 已知1NF關聯表T(a,b,c,d,e,f)有下列FD

◆ FD1 : $a,b \rightarrow c,d,e,f$

◆ FD2 : $a \rightarrow c$

◆ FD3 : $b \rightarrow d$

◆ FD4 : $e \rightarrow f$

◆ 第一步

◆ a

◆ b

◆ a,b

第二步

a,c

b,d

a,b,e,f

第三步

T1(a,c)

T2(b,d)

T3(a,b,e,f)

5-3-3 第三正規化型式 – 3NF(定義)

- ◆ 關聯表R符合第三正規化型式（3NF）
 - ◆ 關聯表符合2NF
 - ◆ 沒有遞移相依的情況發生



5-3-3 第三正規化型式 – 3NF(範例)

- ◆ 例如Instructors關聯表需要第三正規化，因為有遞移相依情況

補充：2NF轉3NF三步驟

- ◆ 檢查每個2NF關聯表綱要，若存在一個遞移相依，將相關的欄位複製到新的一列，視為一個新關聯表綱要。
- ◆ 將該遞移相依中的相依於其他欄位的欄位刪除。
- ◆ 為新關聯表綱要，取適當名稱，並標示其PK。

補充：2NF轉3NF例子

◆ 已知2NF關聯表T3(a,b,e,f)有遞移相依 $e \rightarrow f$

◆ 第一步	第二步	第三步
◆ T1(<u>a</u> ,c)	T1(<u>a</u> ,c)	T1(<u>a</u> ,c)
◆ T2(<u>b</u> ,d)	T2(<u>b</u> ,d)	T2(<u>b</u> ,d)
◆ T3(<u>a</u> ,b,e,f)	T3(<u>a</u> ,b,e)	T3(<u>a</u> ,b,e)
◆ e,f	e,f	T4(<u>e</u> ,f)

5-3-4 Boyce-Codd正規化型式 (BCNF)

- ◆ 如果關聯表至少有二個以上的候選鍵，而且這兩個候選鍵有下列特性，就可能要做BCNF
 - ◆ 兩個都是複合候選鍵
 - ◆ 複合候選鍵之間有共同屬性存在，即至少有一個相同的屬性
- ◆ 如果關聯表沒有上述情況，3NF就等於BCNF

5-3-4 Boyce-Codd正規化型式－範例

- ◆ Students關聯表含學號(sid)、身份證字號(SSN)、課程代號(c_no)與成績(grade)
 - ◆ 有兩個候選鍵(sid, c_no) 和(SSN, c_no)
 - ◆ 候選鍵之間擁有功能相依 $SSN \rightarrow sid$ ，即SSN屬性可以決定sid屬性
 - ◆ $\{ sid, c_no \} \rightarrow SSN$ $SSN \rightarrow sid$
 - ◆ SSN只是候選鍵的一部分，且不是超鍵(Superkey)，違反BCNF

5-3-4 Boyce-Codd正規化型式 – 定義

- ◆ 關聯表R符合BCNF正規化型式
 - ◆ 所有主要的功能相依 $A \rightarrow B$ ，A一定且只可以是候選鍵
 - ◆ 刪除只功能相依候選鍵部分屬性的功能相依

5-3-4 Boyce-Codd正規化型式 – 正規化結果

- ◆ 關聯表Students關聯表的BCNF正規化



補充：3NF轉BCNF二步驟

- ◆ 將3NF關聯表綱要中的特殊相依式子，所牽涉的兩方相關欄位的位置互相調換，使3NF關聯表衰退為1NF關聯表。
- ◆ 利用之前方法，將該1NF關聯表轉換為2NF關聯表，再轉換為3NF關聯表後，自然升級為BCNF關聯表。

補充：3NF轉BCNF例子

◆ 已知3NF關聯表T(a,b,c,d)有下列FD

◆ FD1 : $a,b \rightarrow c,d$

◆ FD2 : $c \rightarrow b$

特殊相依式子

◆ 第一步

◆ T(a,c,b,d)

◆

第二步

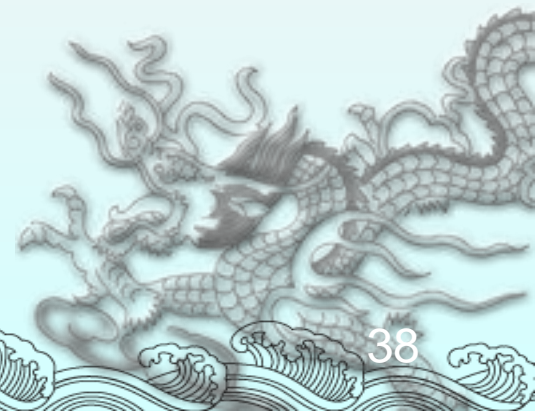
T1(a,c,d)

T2(c,b)

5-4 多重值相依與第四正規化型式

◆ 第四正規化型式

- ◆ 基礎是多重值相依(Multi-valued Dependency, MVD)
- ◆ 屬於功能相依的推廣定義
- ◆ 或是說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



5-5 合併相依與第五正規化型式

◆ 第五正規化型式

- ◆ 基礎是擴充多重值相依的合併相依(Join Dependency)
- ◆ 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