

# 第 5 章

## 關聯式資料模型與關聯式資料庫的限制

### 本章大綱

- 關聯式模型的觀念
- 關聯式限制與關聯式資料庫綱要
- 更新運算動作與處理違反限制的情況

## 學習重點

---

- 關聯式模型的觀念
- 非正式的定義
- 正式的定義
- 定義總整理
- 關聯的特徵
- 關聯式模型的限制

## 學習重點

---

- 鍵值限制
- 超鍵、鍵值、候選鍵與主鍵
- 參考完整性
- 參考完整性限制
- 其他類型的限制
- 關聯的更新運算動作

# 關聯式模型的觀念

---

- 資料的關聯式模型是以「關聯」(Relation) 的觀念為基礎
- 關聯是一種數學觀念，其理論基礎是集合論
- 此模型是在1970年由IBM的E.F. Codd博士的論文所提出

# 關聯 (RELATION) - 非正式的定義

---

- 由資料值所組成的表格
- 關聯可視為是一組橫列 (row) 的集合
- 關聯也可以視為是一組直欄 (**column**) 的集合
- 每個橫列是對應到某個真實世界的實體或關係
- 每個直欄通常是以它的欄位名稱、欄位標題或屬性名稱來稱呼

## 正式的定義 (1/5)

- 一個關聯可以用多種方式來定義
- 關聯的綱要： $R(A_1, A_2, \dots, A_n)$   
關聯網要 $R$ 是由 $A_1, A_2, \dots, A_n$ 這些屬性所定義而成  
例如 -  
    CUSTOMER (Cust-id, Cust-name, Address, Phone#)  
這裡的CUSTOMER是一個用Cust-id、Cust-name、Address、Phone#這4個屬性所定義而成的關聯，每個屬性都有其定義域，也就是一組有效值的集合。例如Cust-id的定義域是6位數的數字

7

## 正式的定義 (2/5)

- 值組 (**tuple**) 是一組值的有序集合 (ordered set)
- 每個值都是從適當的定義域中取得
- 在CUSTOMER表格中的每一列可以稱作一筆值組，其中含有4個值  
例如  $\langle 632895, \text{"John Smith"}, \text{"101 Main St. Atlanta, GA 30332"}, \text{"(404) 894-2000"} \rangle$   
是一筆屬於CUSTOMER關聯的值組
- 一個關聯可以稱作一組值組 (row) 的集合
- 表格中的欄位則稱作關聯的屬性

8

## 正式的定義 (3/5)

- 定義域是具有邏輯上的定義
  - 例如 “USA\_phone\_numbers” 是美國境內所有有效的10位數電話號碼的集合
- 定義域可被指定某種資料型態或資料格式
  - 例如USA\_phone\_numbers的格式可能是：(ddd)-ddd-dddd，其中的d是代表一個10進位數字
  - 日期的格式有很多種，例如yyyy-mm-dd或dd mm,yyyy
- 屬性則是指定定義域所扮演的角色，例如定義域Date可能可以用來定義屬性 “Invoice-date” 和 “Payment-date”

9

## 正式的定義 (4/5)

- 關聯是由定義域的卡氏積 (cartesian product) 所形成的，卡氏積會產生由定義域所衍生出來所有可能值的組合
  - 例如屬性Cust-name的定義域是25個字元的字串，而這些字串在CUSTOMER關聯中所扮演的角色是顧客的姓名
- 正式的說：假定  $R(A_1, A_2, \dots, A_n)$
- $$r(R) \subset \text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n)$$
- R：關聯的綱要
  - r：R的某個特定值

10

## 正式的定義 (5/5)

- Let  $S1 = \{0,1\}$
- Let  $S2 = \{a,b,c\}$
- Let  $R \subset S1 \times S2$

因此，例如  $r(R) = \{ \langle 0,a \rangle, \langle 0,b \rangle, \langle 1,c \rangle \}$  便是關聯  $R$  的某個可能狀態，它是定義在定義域  $S1$  和  $S2$  上，它有 3 筆值組

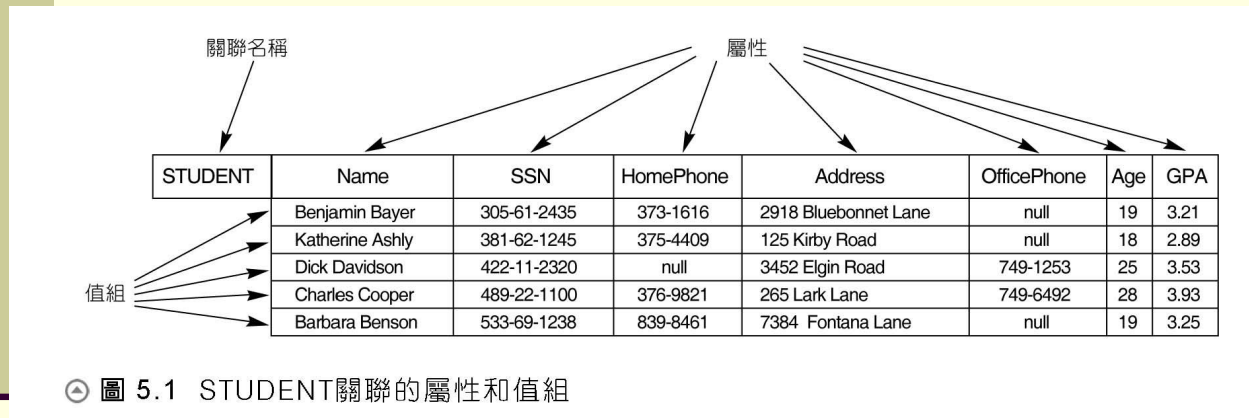
11

## 定義總整理

<u>非正式的術語</u>		<u>正式術語</u>
表格		關聯
欄位 (直欄)		屬性 / 定義域
列 (橫列)		值組
欄位中的值		定義域
表格定義		關聯的綱要

12

## 範例 – 圖5.1



13

## 關聯的特徵

- 關聯  $r(R)$  中的值組順序：關聯中的值組是沒有任何特定順序的
- 關聯網要  $R$  中屬性的順序 (以及值組中資料值的順序)：通常定義在  $R(A_1, A_2, \dots, A_n)$  中的屬性，還有在  $t = \langle v_1, v_2, \dots, v_n \rangle$  裡的值是有順序的 (*ordered*)  
(不過有其他更一般化的關聯定義，這部份是不需要限制順序的)

14

## 關聯的特徵

- 值組中的資料值和空值：在關聯模型中，值組內的每個資料值都是不可分割的 (*atomic*)。而空值 (**null**) 是用來表示屬性值是未知的，或者是不適用於該值組
- 表示法：
  - 將值組  $t$  的成分值以  $t[A_i] = v_i$  來表示 (值組  $t$  的屬性  $A_i$  的值)
  - 同樣的， $t[A_u, A_v, \dots, A_w]$  是代表  $t$  的子值組，裡面分別包括屬性  $A_u, A_v, \dots, A_w$  的值

15

## 關聯的特徵 — 圖 5.2

STUDENT	Name	SSN	HomePhone	Address	OfficePhone	Age	GPA
	Dick Davidson	422-11-2320	null	3452 Elgin Road	749-1253	25	3.53
	Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	null	19	3.25
	Charles Cooper	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
	Katherine Ashly	381-62-1245	375-4409	125 Kirby Road	null	18	2.89
	Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	null	19	3.21

16



# 關聯式模型的限制

- 所謂的限制 (constraint) 是指在所有有效的關聯實例上都要成立的那些條件 (condition)。這裡的限制主要分成3種：
  1. 鍵值限制
  2. 實體完整性限制
  3. 參考完整性限制

17

## 鍵值限制

- R的超鍵 (superkey)：這是R的一組屬性SK的集合，它使得在任何有效的關聯實例 $r(R)$ 中，不會有兩筆值組的SK值完全相同。也就是說，對於 $r(R)$ 中的任兩筆不同的值組 $t_1$ 和 $t_2$ ， $t_1[SK] \neq t_2[SK]$  必須成立

18

## 鍵值

- R的鍵值 (key)：一種「最小的」超鍵，條件是在這個超鍵K中移除了任何屬性後，所形成的新屬性集合就不再是個超鍵

範例：CAR關聯網要：

CAR(State, Reg#, SerialNo, Make, Model, Year)

有兩個鍵值Key1 = {State, Reg#} 和Key2 = {SerialNo}，它們也都是超鍵。而 {SerialNo, Make} 是個超鍵但不是個鍵值

## 候選鍵與主鍵

- 一個關聯網要可能有一個以上的鍵值。這種情況下的每個鍵值稱為候選鍵 (candidate key)
- 假如關聯有數個候選鍵，則可任選其中一個成為主鍵 (primary key)。主鍵的屬性會加上底線

# 鍵值限制範例

CAR	<u>LicenseNumber</u>	EngineSerialNumber	Make	Model	Year
	Texas ABC-739	A69352	Ford	Mustang	96
	Florida TVP-347	B43696	Oldsmobile	Cutlass	99
	New York MPO-22	X83554	Oldsmobile	Delta	95
	California 432-TFY	C43742	Mercedes	190-D	93
	California RSK-629	Y82935	Toyota	Camry	98
	Texas RSK-629	U028365	Jaguar	XJS	98

圖 5.4 CAR關聯有兩個候選鍵：LicenseNumber與EngineSerialNumber

21

# 實體完整性

- 關聯式資料庫綱要 (relational database schema) : 一組屬於同一資料庫的關聯網要的集合S，S是該資料庫的名稱

$$S = \{R_1, R_2, \dots, R_n\}$$

- 實體完整性 (Entity Integrity) : S中的每個關聯網要R的主鍵屬性PK，在r(R)中的任何值組都不能有空值。這是因為主鍵值是用來識別 (identify) 個別的值組

對於在r(R)中的任何值組t，t[PK] ≠ null

注意：R的其他屬性也可能會限制不允許是空值，即使它們不一定是主鍵的成員

22

# 參考完整性

- 這個限制牽涉到兩個關聯 (之前的限制只牽涉一個關聯)
- 使用指定兩個關聯的值組彼此的關係，也就是參考關聯 (referencing relation) 與被參考關聯 (referenced relation)
- 在參考關聯 $R_1$  中的值組具有屬性FK (稱作外來鍵屬性)，它會參考到被參考關聯 $R_2$  的主鍵屬性PK。因此  
假如  $t_1[FK] = t_2[PK]$ ，則稱 $R_1$  中的值組 $t_1$  參考 $R_2$  中的值組 $t_2$
- 參考完整性限制的圖形表示法，是從 $R_1.FK$ 畫出一條有箭頭的弧線到 $R_2$ 上

23

# 參考完整性限制

## 此限制的正式定義

在參考關聯 $R_1$  外來鍵欄位FK的值可以是以下其中之一：

- (1) 在被參考關聯 $R_2$  中相對應主鍵的某個已存在的主鍵值，或者...
- (2) 是個空值

在 (2) 情況下， $R_1$  的FK不應該是本身主鍵的一部分

24

# 其他類型的限制

## 語意完整性限制 (Semantic Integrity Constraint)

- 是根據應用本身的語意，無法由模型本身來表達
- 例如「員工在所有計畫上工作的每週最大時數為56」就算是這類限制
- 這類限制是利用一種限制指定語言 (constraint specification language) 來表達
- SQL-99中有觸發 (trigger) 和宣示敘述可用來表達

25

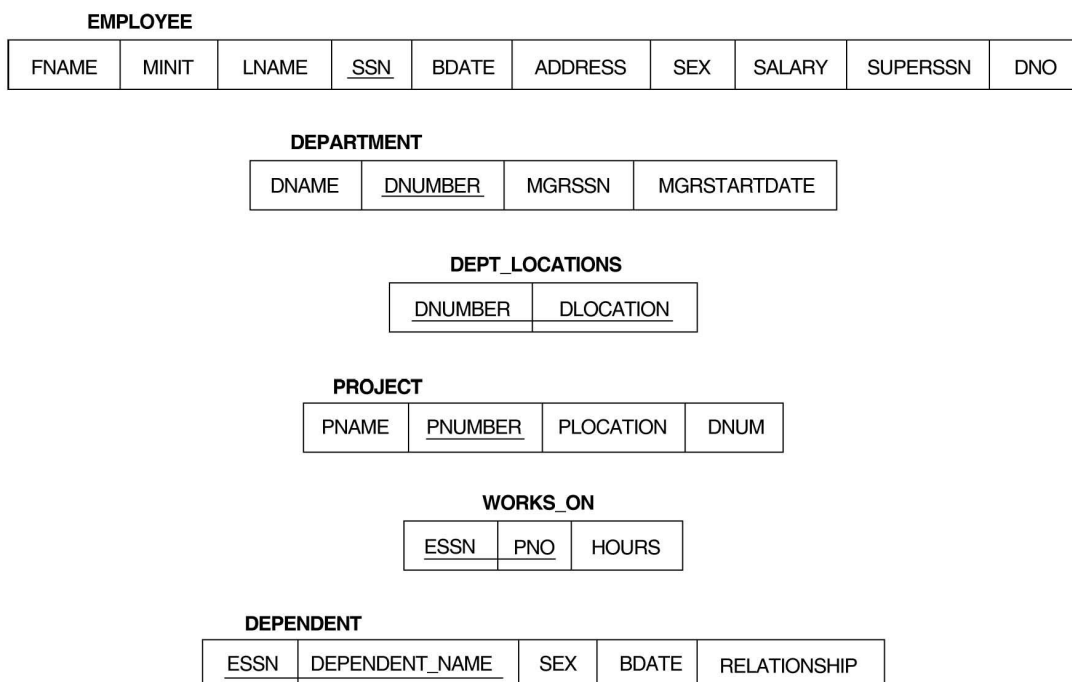


圖 5.5 COMPANY 關聯式資料庫綱要的綱要圖

26

EMPLOYEE	FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5	
Franklin	T	Wong	333445555	1965-12-08	638 Voss, Houston, TX	M	40000	888665555	5	
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4	
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4	
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5	
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5	
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4	
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	null	1	

DEPT_LOCATIONS	DNUMBER	DLOCATION
	1	Houston
	4	Stafford
	5	Bellaire
	5	Sugarland
	5	Houston

DEPARTMENT	DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research		5	333445555	1988-05-22
Administration		4	987654321	1995-01-01
Headquarters		1	888665555	1981-06-19

WORKS_ON	ESSN	PNO	HOURS
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0
	453453453	2	20.0
	333445555	2	10.0
	333445555	3	10.0
	333445555	10	10.0
	333445555	20	10.0
	999887777	30	30.0
	999887777	10	10.0
	987987987	10	35.0
	987987987	30	5.0
	987654321	30	20.0
	987654321	20	15.0
	888665555	20	null

PROJECT	PNAME	PNUMBER	PLOCATION	DNUM
	ProductX	1	Bellaire	5
	ProductY	2	Sugarland	5
	ProductZ	3	Houston	5
	Computerization	10	Stafford	4
	Reorganization	20	Houston	1
	Newbenefits	30	Stafford	4

DEPENDENT	ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
	333445555	Alice	F	1986-04-05	DAUGHTER
	333445555	Theodore	M	1983-10-25	SON
	333445555	Joy	F	1958-05-03	SPOUSE
	987654321	Abner	M	1942-02-28	SPOUSE
	123456789	Michael	M	1988-01-04	SON
	123456789	Alice	F	1988-12-30	DAUGHTER
	123456789	Elizabeth	F	1967-05-05	SPOUSE

圖 5.6 COMPANY的關聯式資料庫的其中一個可能的資料庫狀態

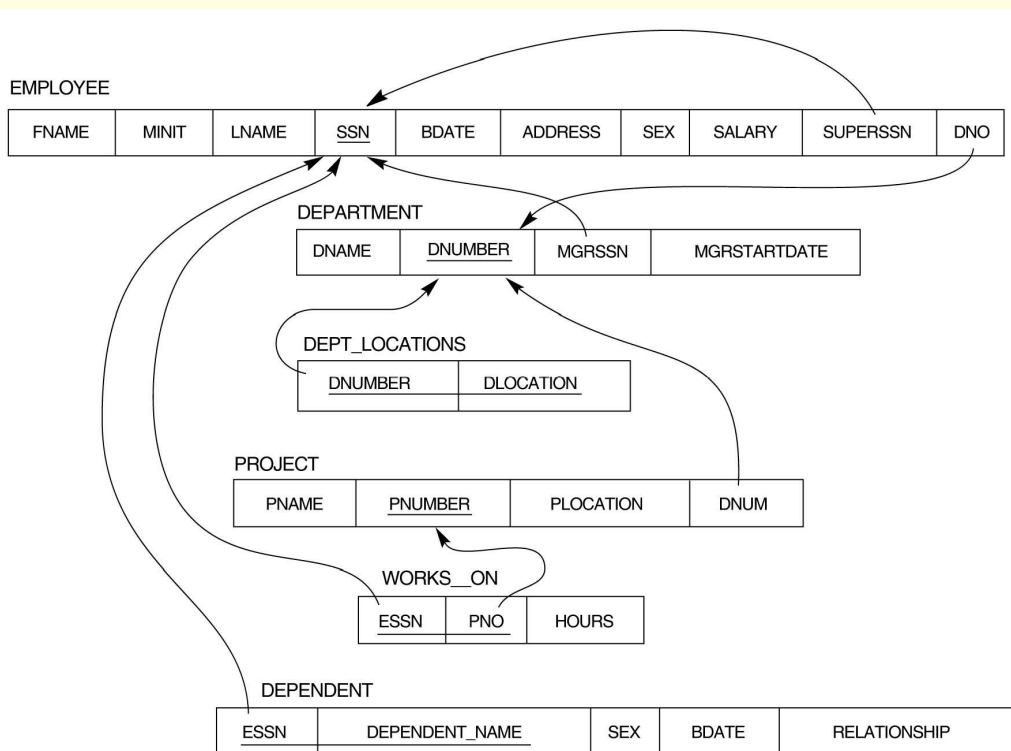


圖 5.7 顯示COMPANY關聯式資料庫綱要的參考完整性限制

## 關聯的更新運算動作

- 插入 (INSERT) 一筆值組
- 刪除 (DELETE) 一筆值組
- 修改 (MODIFY) 一筆值組
  - 在進行更新運算時，原本指定的完整性限制仍然不應該違反
  - 有時可能需要將數個更新運算組合在一起
  - 更新運算可能會傳播而自動引發其他的更新運算。為了維持完整性限制可能必須如此

29

## 關聯的更新運算動作

- 假如發生違反完整性限制的情況，可以採取以下這些措施：
  - 取消會導致這個違反情況的運算 (REJECT 選項)
  - 繼續進行但通知使用者有發生違反情況
  - 觸發額外的更新動作，來修正這個違反情況 (CASCADE 選項、SET NULL 選項)
  - 執行由使用者所定義的錯誤修正程序

30

## 隨堂練習

下列是某個資料庫的幾個關聯，裡面記錄各種課程的學生修課人數，以及每一門課所採用的教科書：

STUDENT(SSN, Name, Major, Bdate)

COURSE(Course#, Cname, Dept)

ENROLL(SSN, Course#, Quarter, Grade)

BOOK\_ADOPTION(Course#, Quarter, Book\_ISBN)

TEXT(Book\_ISBN, Book\_Title, Publisher, Author)

請畫出這個綱要圖，並畫出外來鍵。(取自習題5.8)

31

## 學習評量

- 為何關聯中的值組沒有順序？
- 為何關聯中不允許重複值組的存在？
- 鍵值與超鍵之間有何不同？
- 請討論在關聯中會導致空值出現的各種理由。
- 請說明實體完整性與參考完整性限制，並敘述它們的重要性。

32