
Ch3

關聯式資料庫模型

資料模型(式)

- 既有(舊)資料庫模型(Legacy database model)
 - 階層式模型(Hierarchical model)
 - 網狀式模型(Network model)
- 關聯式模型 (Relational model)
 - 目前主流資料模式
 - 最早可用的商業版本出現在1980年代早期
 - 已經在許多大型商業系統上實作
- NoSQL
 - MongoDB, Firebase,

☆要轉換到關聯怎麼做?
用E/R圖來轉換
舊系統→E/R圖
→關聯資料庫

關聯式模型的觀念

- 資料庫是由許多關聯 (relation)(資料表) 所組成的集合
- 資料表(table)
 - 列 (row)
 - 代表相關資料值的集合
 - 對應到真實世界中的某個實體或關係
 - 又稱值組 (tuple)(學妹行上)
 - 欄位名稱
 - 說明每一列的各屬性 (attribute) 所代表的意義

學術上
Relation

商業上
Table

一筆資料:

學術 Tuple

商業: row, record,
Entry Record

學術: Attribute

商業: Column

傳統檔案: Field

關聯式模型的觀念 (續)

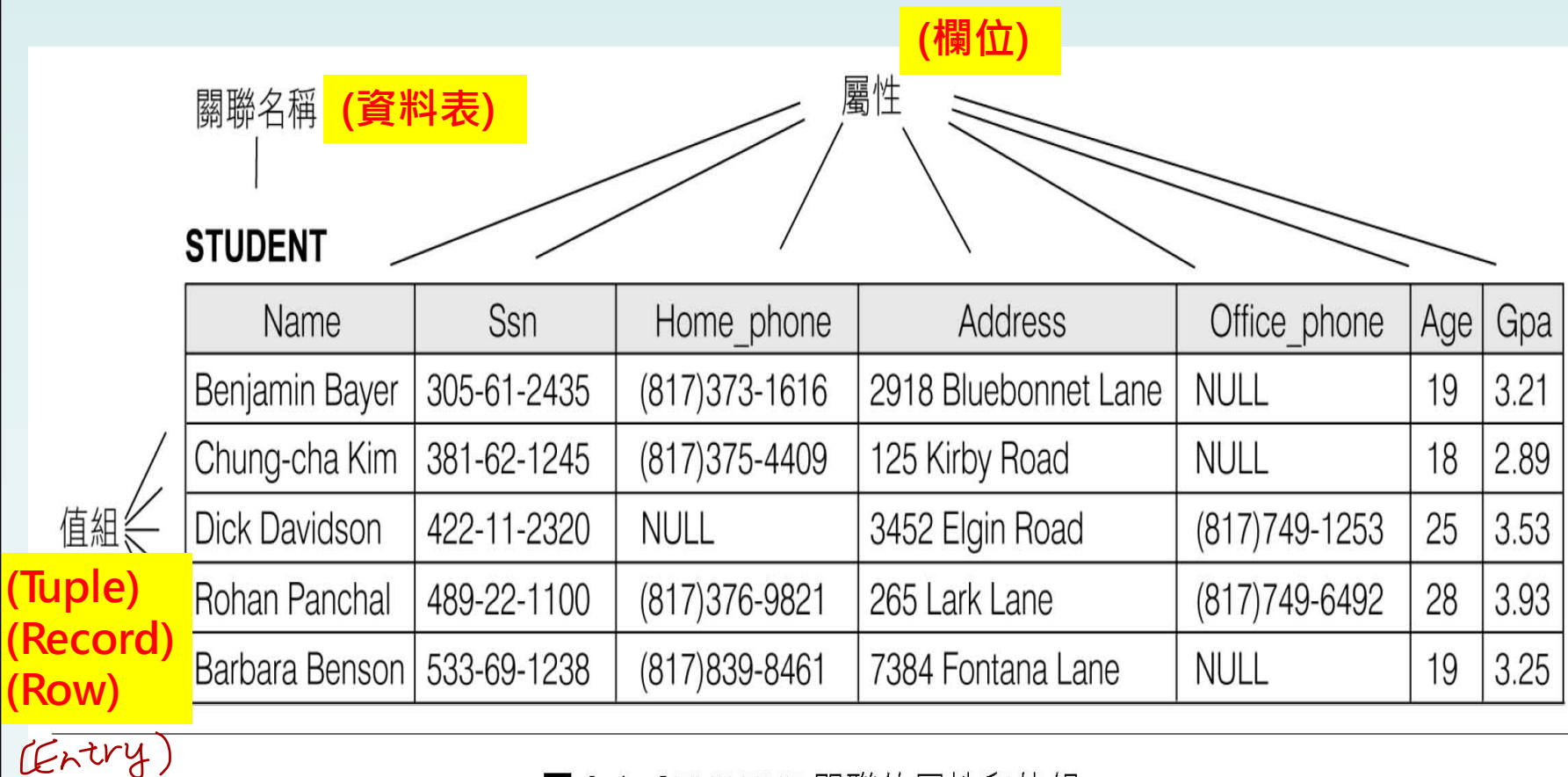


圖 3.1 STUDENT 關聯的屬性和值組

定義域、屬性、值組與關聯

- 定義域 (domain, D): 就是欄位的**有效範圍**
 - 由一組不可分割的單元值所組成的集合
- 單元 (atomic) 值
 - 每個值是**不可分割的**
 - 不可以是**多重值**(multivalue)
 - 不可以是**重複群組**(repeating group)
↳ 單獨拉出來, 跟多重值方法相近, 見 D04
- 指定定義域
 - 常用的方法是替每個定義域指定**資料型態**
 - Oracle可以設定某個名稱的定義域(Domain), 讓欄位套用

定義域、屬性、值組與關聯 (續)

■ 關聯網要(結構)(Schema) R

- 以 $R(A_1, A_2, \dots, A_n)$ 來表示
- 由關聯名稱 R 與一串屬性 $A_1, A_2, \dots, A_n$ (Attributes) 所組成

$R(A_1, A_2)$ 則
維度為 4

■ 關聯的 **維度** (degree或arity)

- 關聯網要(結構)的 **屬性個數** n 稱為維度

■ **基數**(資料筆數)(cardinality)

- 一個定義域內資料的總筆數 有幾筆資料

定義域、屬性、值組與關聯 (續)

■ 關聯 (或關聯狀態) $r(R)$

- 關聯 $r(R)$ 是定義於 $\text{dom}(A_1), \text{dom}(A_2), \dots, \text{dom}(A_n)$ 的 n 階數學關聯
- R 為定義域卡氏積 (Cartesian product) 的子集之一
 - $r(R) \subseteq (\text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n))$

■ 關聯 (或關聯狀態)

- **N**值組 (n-tuple) 所組成的集合 $r = \{t_1, t_2, \dots, t_n\}$
- 每個 n 值組 t
 - 是由 n 個 $t = \langle v_1, v_2, \dots, v_n \rangle$ 組成的有序串列
 - 其中每個 $v_i (1 \leq i \leq n)$ 都是 $\text{dom}(A_i)$ 的一個元素或是個特殊的 NULL 值

關聯的特徵

- 關聯中的值組順序
 - 關聯定義為一群相關值組的集合
 - 關聯中值組的順序不重要
 - 元素(屬性)順序不重要

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21

圖 3.2 圖 3.1 的 STUDENT 關聯，但以不同的順序來排列值組

關聯的特徵 (續)

- 值組中的資料值和NULL
 - 不能多值
/ 不能重複群組
 - 值組內的每個資料值都是單元值(atomic)
 - 關聯式模型
 - 重複群組屬性與多重值屬性都是不被允許的
 - 第一正規化 (first normal form) 的定義
 - 多重值屬性
 - 必須以另外的關聯來表示
 - 複合屬性(重複群組屬性)
 - 必須以另外的關聯來表示

關聯的特徵 (續)

■ NULL值

- 代表某些屬性的值可能是未知的，或者是不適用於該值組
- NULL值可能有好幾種意義
 - 未知的資料值
 - 資料值無法取得
 - 不適用於此值組的屬性 (也稱作「資料值未定義」)

關聯式模型的習慣符號表示法

- N維度的關聯綱要(結構) R
 - 表示成 $R(A_1, A_2, \dots, A_n)$
- 大寫字母 Q 、 R 、 S
 - 代表關聯名稱
- 小寫字母 q 、 r 、 s
 - 代表關聯狀態(某特定時間點)
- 字母 t 、 u 、 v
 - 代表值組(tuple)

關聯式模型的符號表示法

- 關聯網要(結構)的名稱：STUDENT
 - 代表目前在該關聯中的值組集合
- 表示法：STUDENT(Name, Ssn, ...)
 - 只代表關聯網要(結構)
- 屬性名稱 A 可用其所屬的關聯名稱 R 來加以指定
 - $R.A$

關聯式模型的符號表示法

- 在關聯 $r(R)$ 中的 n 值組 t
 - 表示成 $t = \langle v_1, v_2, \dots, v_n \rangle$
 - 其中 v_i 是對應到屬性 A_i
- 值組的成分值：
 - $t[A_i]$ 與 $t.A_i$ 都代表在 t 中屬性 A_i 的資料值 v_i
 - $t[A_{u_i}, A_{w_i}, \dots, A_{z_i}]$ 與 $t.(A_{u_i}, A_{w_i}, \dots, A_{z_i})$ 都是代表從 t 對映到此串列所指定屬性值的 $\langle v_{u_i}, v_{w_i}, \dots, v_{z_i} \rangle$ 子值組

關聯式模型的限制

■ 限制 (constraint)

- 在資料庫中針對實際存放的資料(值)的限制(規則)

■ 限制種類

- 定義域限制

1. 資料不重複 2. Not Null

- 鍵值限制與空值限制

- 實體完整性限制

- 參考性整合限制 → 由key與fk建構出來的條件 (cascade / Restricted)

- 企業規則限制

定義域限制

■ 一般包括

■ 資料型態

- 數值：如整數、實數；字元、布林值、日期、時間、時間戳記、貨幣、其他特殊的資料型態
- 固定長度的字串、變動長度的字串
char *varchar*

■ 是否允許空值(null)

■ 是否具唯一性 *⇒ 不一定key*

鍵值限制與空值限制

- 資料庫內沒有任何兩個值組內的全部屬性值完全相同
- 超鍵(super key, SK)
 - 在關聯R中，包含key的屬性(欄位) 集合
- 鍵值(key)
 - R的超鍵 之一
 - 從K移除任何一個屬性A之後，所形成的新集合K'不再是R的一個超鍵(最小的超鍵)

只要有包含 key 就算
如 $R(\underline{a}, \underline{b}, c, d)$ $a, b \Rightarrow SK$
 $b, c \Rightarrow SK$
 $c, d \Rightarrow \cancel{SK}$

鍵值限制與空值限制 (續)

- 鍵值(key value)會滿足以下兩個條件：
 - 在同一個關聯中，二筆不同的值組不能有完全相同的鍵值(唯一性)
 - 最小的超鍵
 - 如果移除了任何屬性，就無法保有第一個條件的唯一性限制

鍵值限制與空值限制 (續)

- 候選鍵(Candidate key)、關聯的主鍵(或主索引) (Primary key)
 - 關聯綱要(結構)可能有一個以上的鍵值，**只能一個選為主鍵**，其餘均為候選鍵
 - 其他候選鍵**亦為唯一鍵值** *一個檔只有一個主索引*
主索引可包含好幾個欄位

關聯式資料庫與關聯式資料庫綱要(結構)

■ 關聯式資料庫綱要(結構) S

很多個資料表

- 一組關聯綱要(結構)的集合 $S = \{R_1, R_2, \dots, R_m\}$
- 加上一組完整性限制集合 IC(Integrity constraint)

■ 關聯式資料庫狀態

relation 資料表

- 關聯狀態的集合 $DB = \{r_1, r_2, \dots, r_m\}$
- 其中每一個 r_i 都是 R_i 的狀態，並且 r_i 關聯狀態會滿足 IC 中所指定的完整性限制

關聯式資料庫與關聯式資料庫綱要(結構) (續)

- 無效狀態
 - 沒有遵守所有完整性限制
- 有效狀態
 - 滿足IC中所有的限制

實體完整性、參考完整性與外來鍵

- **實體完整性限制(entity constraint)**
 - 主鍵的值不可以等於NULL
- **參考完整性限制(referential integrity constraint)**
 - 指定在兩個關聯之間，以主索引與外來鍵所建立的關係，以保持兩個關聯值組間的一致性

實體完整性、參考完整性與外來鍵 (續)

- 外來鍵規則(Foreign key)
 - FK與擁有者實體之主鍵屬性PK有相同的定義域，並由此PK-FK形成父子關係
- 參考完整性限制的圖形表示法
 - 從外來鍵畫一條有箭頭的弧線到它所參考關聯的欄位上
- 在關聯式資料庫綱要(結構)中應該指定所有的完整性限制

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
-------	----------------	---------	----------------

DEPT_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

WORKS_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
-------------	-----------------------	-----	-------	--------------

圖 3.7 顯示 COMPANY 關聯式資料庫綱要的參考完整性限制

圖 3.6 COMPANY 關聯式資料庫綱要的其中一個可能的資料庫狀態

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	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

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

圖 3.6 COMPANY 關聯式資料庫綱要的其中一個可能的資料庫狀態

WORKS_ON

<u>Essn</u>	<u>Pno</u>	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

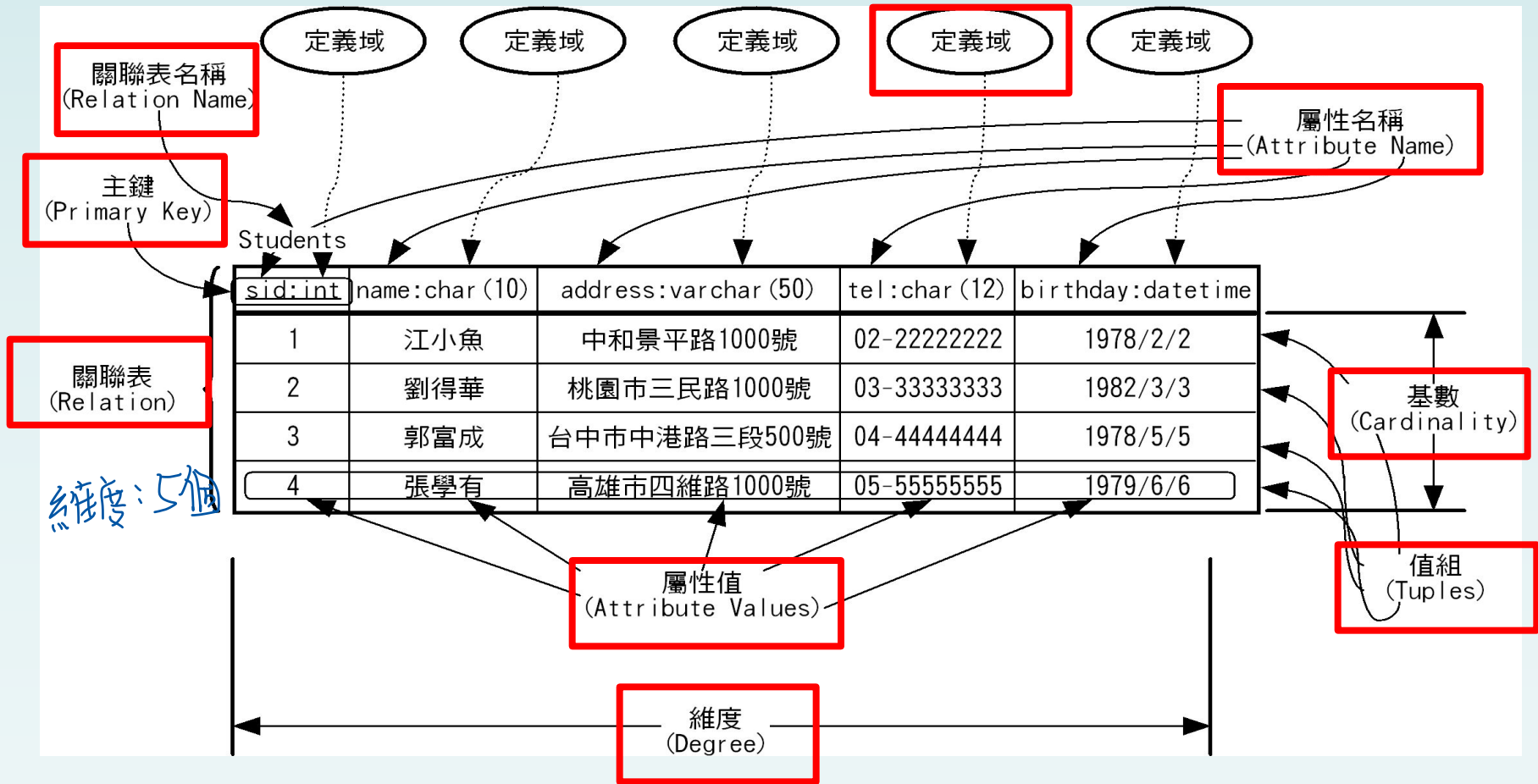
PROJECT

<u>Pname</u>	<u>Pnumber</u>	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

<u>Essn</u>	<u>Dependent_name</u>	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

關聯表綱要-相關術語



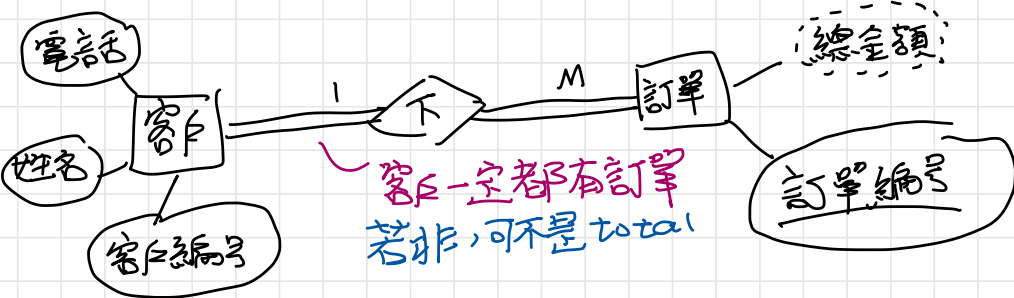
關聯式資料庫模型	資料庫	檔案系統
關聯表 (Relation)	資料表 (Table)	檔案 (File)
值組 (Tuple)	資料列 (Row)	記錄 (Record)
屬性 (Attribute)	資料欄 (Column)	欄位 (Field)

2個是獨立的
子檔不受影響，易造錯誤，除非2者獨立

參考性完整限制

	刪除有相關子檔 的父檔資料	更新有相關子檔 的父檔之主索引	新增資料到 無父檔資料之子檔
忽略 (Ignore)	可以	可以	可以
限制型 (Restricted)	不允許	不允許	不允許
連鎖反應型 (Cascade)	相關子檔資料 一併刪除	相關子檔資料 之外來鍵一併更新	不允許
設為預設值 (set Default)	相關子檔資料之外來 鍵一併更新為預設值	相關子檔資料之外來 鍵一併更新為預設值	子檔資料之外來鍵 為預設值
設為虛值 (set Nullifies)	相關子檔資料之外來 鍵一併更新為虛值	相關子檔資料之外來 鍵一併更新為虛值	子檔資料之外來鍵 為虛值

- 一般新增先父後子，刪除先子後父（操作範例）



Restricted:

1-2-4 不能更改
或刪除。(有子檔)
也不能新增 09
可以

父子關係 例子

	FK		
1	-	-	
2	-	-	
3	-	-	
4	-	-	

01	1	
02	1	
03	2	
04	2	
05	2	
06	3	
07	3	
08	3	

Cascade:

1-2-4 刪除後, 子檔也會刪。
改 ID, 子檔也會改。
不能新增 09

set Default

1 若刪除, 則 FK FK 就改為預設值
若更改主索引, 則改為預設值
新增 09, 則設 FK 為預設值

set Nullifies

1 若刪除, 則 FK FK 就改為 null
若更改主索引, 則改為 null
新增 09, 則設 FK 為 null

以SQL設定參考性整合

•SQL/92、SQL/99 支援全部5種狀況

- 預設值是限制型(**NO ACTION**) (不允許 *delete/update*)
- 連鎖反應型(**CASCADE**) : 相關資料一併刪除、更新
- 相關外來鍵值設為虛值或預設值(**SET NULL / SET DEFAULT**)

```
CREATE TABLE Enrolled
(sid CHAR(20),
cid CHAR(20),
grade CHAR(2),
PRIMARY KEY (sid,cid),
FOREIGN KEY (sid)
REFERENCES Students
ON DELETE CASCADE
ON UPDATE SET
DEFAULT )
```