第八章 正規化

- 好的關聯綱目特性
 - ∞ 意義明確

 - (2) 資料不重複儲存(2) 屬性值很少有空值(2) 關聯的JOIN不會產生虛假序列值
- 函數相依
- 基本正規化

 - (28 <u>第一正規式</u> (28 <u>第二正規式</u> (28 <u>第三正規式</u>
- 進階正規式
 - Boyce-Codd正規式 第四正規式

 - ∞ 第五正規式

好的關聯綱目特性

- 意義明確
 - 記載迷你世界裡的同一類實體或關係
 - 以下的關聯包括現實世界裡的交易和會員實體,所以意義不明確

TransactionMember

I	交易編號	交易方式	交易時間	會員編號	會員ID	姓名	生日
	<u>tNo</u>	method	transTime	mld	pld	name	birthday
ı	91100	cart	2005-02-02:18:30:00	a0911234	A122555888	Tony	1980-12-12
ı	92666	cart	2005-10-10:22:10:30	c0927777	B123123123	Su	1982-06-06
ı	91888	fax	2005-09-10:10:10:00	a0910001	A220123456	Jenny	1979-01-01
ı	92333	email	2005-10-15:09:00:00	c0927777	B123123123	Su	1982-06-06
ı	90111	cart	2005-05-05:12:30:30	b0905555	C200456789	Jennifer	1974-03-04
ı	92555	cart	2005-11-11:09:10:00	b0922468	R100200300	Jackson	1980-03-30

好的關聯綱目特性 (Cont.)

- 資料不重複儲存
 - ∞ 空間浪費
 - ∞ 更新異常
- 以下關聯 (OrderProduct)裡「任賢齊專輯三」的資料重複 儲存

OrderProduct	會員編號	購物車產生時間	商品編號	訂購數量	商品名稱	定價	種類
Olderriedae.	mľd	<u>cartTime</u>	pNo	amount	p <u>N</u> ame	unitPrice	catalog
	a0911234	2005-02-02:18:00:30	b30999	1	資料庫理論與實務	500	Book
	b0905555	2005-05-05:12:00:00	v01888	3	哈利波特:混血王子的背叛	450	DVD
	c0927777	2005-10-10:22:00:00	d11222	1	任賢齊專輯三	300	CD
	c0927777	2005-10-10:22:00:00	d20777	1	蔡依林專輯二	350	CD
	c0927777	2005-10-10:22:00:00	v00111	2	英雄	400	DVD
	b0922468	2005-11-11:09:00:30	b10234	5	管理資訊系統概論	600	Book
	b0922468	2005-11-11:09:00:30	b40555	10	系統分析理論與實務	550	Book
	a0010001	2005-00-00-10-00-10	411222	1	仁堅	300	CD

(b)

©黃三益2007 資料庫的核心理論與實務第三版 8-3

好的關聯綱目特性 (Cont.)

- ■有三種更新異常
 - ∞新增異常

INSERT INTO TransactionMember(tNo, mld, name) VALUES ('93000' , 'c0927777' , 'Chen');

∞刪除異常

DELETE TransactionMember WHERE tNo = '90111';

∞修改異常

UPDATE TransactionMember
SET name = 'Huang'
WHERE tNo = '92333;

©黃三益2007 資料庫的核心理論與實務第三版

練習8-1

- 考慮圖8-2(b)的OrderProduct範例關聯,請問刪除哪一筆記錄後會造成刪除異常?
- Ans:

除了第3筆和第8筆記錄外,刪除其他記錄都會造成該商 品資訊消失,而形成刪除異常。

> ©黃三益**2007** 料庫的核心理論與實務第三版

8-5

好的關聯綱目特性 (Cont.)

- 屬性值很少有空值
 - ∝一個設計不良的綱目可能會造成有些屬性有許多空值

Product Author	商品組號	商品名稱	定價	種類	創作者名稱1	創作者名稱2	創作者名稱3	創作者名稱4
Product, Onser	pNo		unitPrice	catalog	authorName1	authorName2	authorName3	authorName4
	b30999	資料庫理論與實務	500	Book	Huang	null	null	null
	d11222	任賢齊專輯三	300	CD	William	null	null	null
	b20666	OLAP進階	500	Book	Sandra	null	null	null
	b10234	管理資訊系統概論	600	Book	Lin	null	null	null
	b40555	系統分析理論與實務	550	Book	Wu	null	null	null
	d20777	蔡依林專輯二	350	CD	Jolin	null	null	null
	v01888	哈利波特:混血王子的背叛	450	DVD	J.K.	null	null	null
	d03333	5566專輯	450	CD	Jackey	David	Tom	null
	b51111	電子商務理論與實務	700	Book	Lai	Hwung	Lin	null
	v00111	拉排	400	DVD	null	null	null	null

⇔缺點

- ■浪費空間
- 彙總函數應用在有空值的屬性時,其意義不明確
- 空值的含義不唯一

©黃三益2007 資料庫的核心理論與實務第三版

好的關聯綱目特性 (Cont.)

- 關聯的JOIN不會產生虛_{Record2} 假序列值
 - □以上三種問題都是因為 一個關聯綱目包含過多 Product2 的屬性
 - 若將一個關聯綱目分解 成數個,分解的方式要 注意

交易編號	商品编號	交易數量	售價	種類
<u>tNo</u>	pNo	amount	salePrice	catalog
91100	b30999	1	450	Book
92555	b10234	5	3000	Book
92666	d11222	1	300	CD
92666	d11222	1	300	CD

商品名稱	定價	種類
<u>pName</u>	unitPrice	catalog
資料庫理論與實務	500	Book
管理資訊系統概論	600	Book
任賢齊專輯三	300	CD

(b)

RecordProduct	交易編號	商品編號	交易數量	售價	商品名稱	定價	種類
Recordi foduci	<u>tNo</u>	<u>pNo</u>	amount	salePrice	p <u>N</u> ame	unitPrice	catalog
	91100	b30999	1	450	資料庫理論與實務	500	Book
	92555	b10234	5	3000	管理資訊系統概論	600	Book
	92666	d11222	1	300	任賢齊專輯三	300	CD

©黃三益**2007** 資料庫的核心理論與實務第三版 8-

<u>交易編號</u> <u>tNo</u>	商品編號 pNo	交易數量 amount	售價 salePrice	商品名稱 pName	定價 unitPrice	種類 catalog
91100	b30999	1	450	資料庫理論與實務	500	Book
91100	b30999	1	450	管理資訊系統概論	600	Book
92555	b10234	5	3000	資料庫理論與實務	500	Book
92555	b10234	5	3000	管理資訊系統概論	600	Book
92666	d11222	1	300	任賢齊專輯三	300	CD

- •上例產生了兩筆虛假序列值(以綠色底色顯示)
- •無損JOIN的分解(Lossless join decomposition):切割之後的兩個關聯,其共同屬性必須是其中一個關聯的關聯鍵
- •以上的分解(Product2和Record2關聯)不滿足無損JOIN的分解,因為共同屬性「catalog」在兩個關聯裡都不是關聯鍵

©黃三益2007 資料庫的核心理論與實務第三版

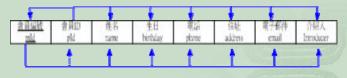
函數相依

- 定義:兩個(群)屬性間存在的一種類似函數 (function)裡定義域(domain)和對應域(codomain)的關係
 - - 每個mld剛好有一組{name, birthday}
 - 不會有一個mld有兩組或以上的{name, birthday}
 - α請注意{name} →{mld}不成立
 - 因為可能有同名同姓的會員
- 對於函數相依X→Y, 我們稱 X決定Y 或Y 函數相依 於X

©黃三益**2007** 資料庫的核心理論與實務第三版 8-9

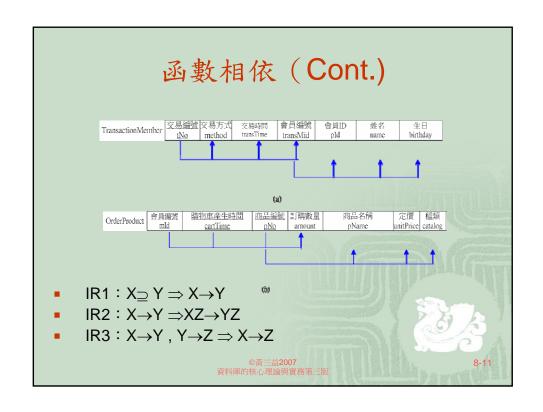
函數相依 (Cont.)

根據定義,一個關聯的任何關聯鍵皆可決定其所有 屬性



- 屬性和屬性間的函數相依是一種**語意**的關係,必須 由人工仔細推敲來訂定
- 但有些函數相依仍然可以有規則可以推論而得

©黃三益2007 資料庫的核心理論與實務第三版



函數相依 (Cont.)

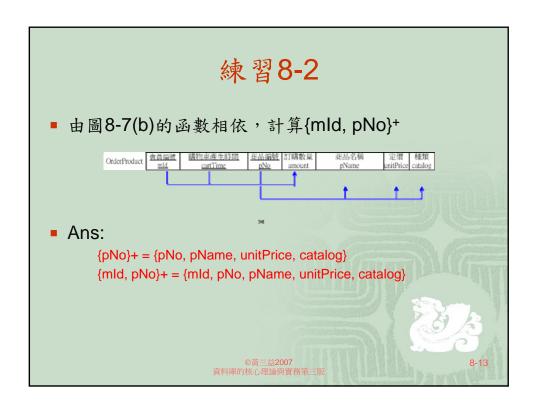
- 一個屬性集合的泛封閉集合:給定一些函數相依, 我們可以推導出這些屬性可以決定的所有屬性,稱 為泛封閉集合
 - ∞ 圖8-7(a)

{tNo}+ = {tNo, transmld, method, transTime, pld, name, birthday} {transMid}+ = {transMid, pld, name, birthday}

∞ 圖8-7(b)

{pNo}+ = {pNo, pName, unitPrice, catalog}
{mId, cartTime, pNo}+ = {mId, cartTime, pNo, amount, pName, unitPrice, catalog}

©黃三益2007 資料庫的核心理論與實務第三版





第一正規式

- R滿足第一正規式(1NF)如果
 - R的每個屬性都是簡單且單值
- 每個合法的關聯綱目都滿足第一正規式
- 如何分解多值屬性
 - ∞ 處理方式一:
 - 維持同樣的屬性,但關聯主鍵要加上該多值屬性。
 - ∞ 處理方式二:
 - 產生另一個關聯,此關聯包括原主鍵和該多值屬性

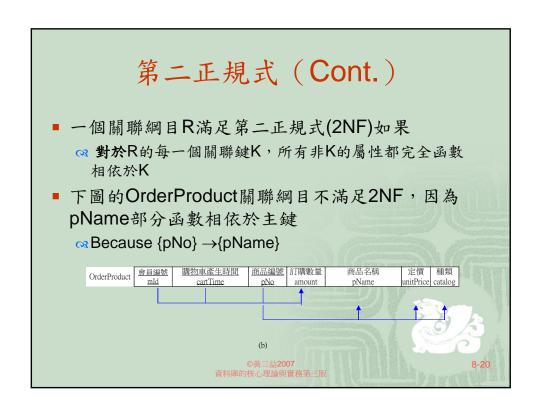
©黃三益**2007** 資料庫的核心理論與實務第三版

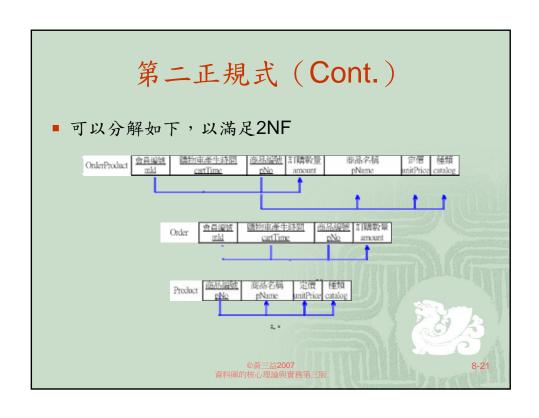
	多	值屬	性	處	足理方	式	_
Framsaction_Products	交易編集 tNo	會員編號 transmid	交易方式	t that	交易時間 transTime		商品 products
	91100	a0911234	cert		2005-02-02:18:30	:00	(b30999)
	92666	c0927777	cent		2005-10-10:22:10	30	{d11222,d20777,v00111}
Ï	92333	c0927777	ameil		2005-10-15:09:00	:00	(b51111)
	91888	e0910001	fax.		2005-09-10:10:10	:00	{b40555, d03333}
	90111	10905555	cent		2005-05-05:12:30	:30	{v01888}
	92555	10922468	cert		2005-11-11:09:10	:00	(b10234, b40555)
Transaction_Product	交易編號 t <u>No</u>	會員編號 transMid	交易方式 method		交易時間 transTime	商品 produc	t to the same of t
	91100	a0911234	cart		-02-02:18:30:00		PARTICIPATION OF THE RESIDENCE OF THE PARTY
	92666 92666	c0927777 c0927777	cart cart		5-10-10:22:10:30 5-10-10:22:10:30		OF NOTE THE REPORT OF THE PART AND ADDRESS OF THE PARTY.
	92666	c0927777	cart		5-10-10:22:10:30		
	92333	c0927777	email	2005	-10-15:09:00:00	b51111	1 337
	91888	a0910001	fax		-09-10:10:10:00		
	91888	a0910001	fax		5-09-10:10:10:00		
	90111	b0905555	cart		-05-05:12:30:30		
	92555	b0922468	cart		5-11-11:09:10:00		The Print Control of the Control of
	92555	b0922468	cart	2005	5- 11 -11:09:10:00	b40555	

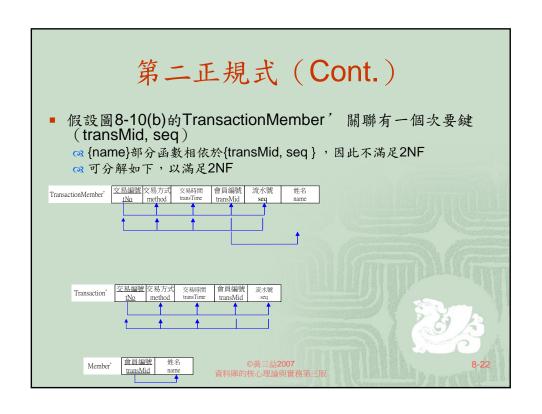












練習8-4

- 考慮<u>圖8-8(b)</u>的Transaction_Product關聯綱目,請問其是否滿足2NF?
- Ans

因為{tNo,product} → { transMid, method, transTime},

且{tNo} → { transMid, method, transTime} 所以, { transMid, method, transTime}為部份函數相依於{tNo, product},並不符合第二正規化

©黃三益**2007** 資料庫的核心理論與實務第三版 8-23

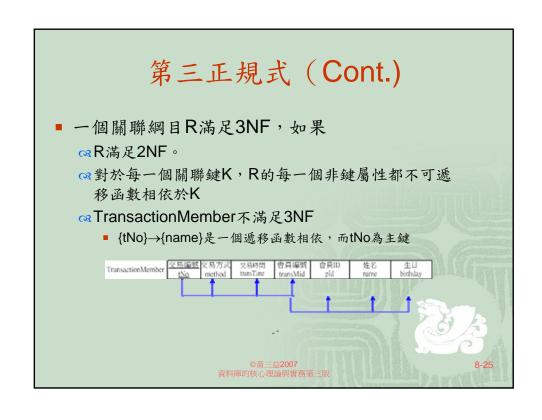
第三正規式

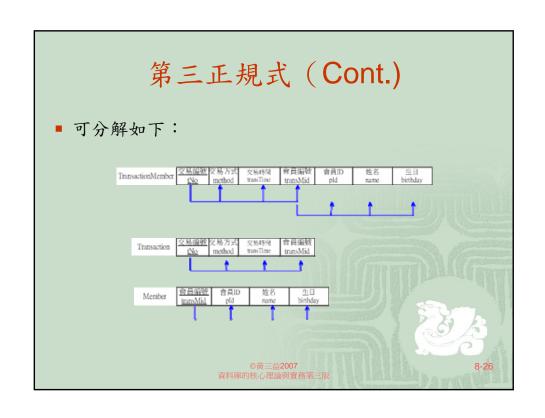
- 第三正規式(簡稱3NF)是根據遞移函數相依的 概念
 - 對於一個函數相依X→Y,如果存在著另外兩個函數相依:X→Z和Z→Y,且Z不為超級鍵,則稱X→Y為遞移函數相依
 - ∝下圖TransactionMember關聯



{tNo}→{name}是一個遞移函數相依
 □ 因為{tNo}→{transMid}且 {transMid}→{name}

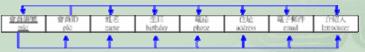
©黃三益2007 資料庫的核心理論與實務第三版





練習8-5

■ 考慮下圖Member關聯的函數相依,有人說此關聯不滿足 3NF,因為 $\{mld\}\rightarrow \{name\}$ 是遞移函數相依,可以由 $\{mld\}\rightarrow \{pld\}\rightarrow \{name\}$ 推導而得。你認為呢?



- Ans:
 - 如上圖函數相依所示,雖然{mld}→{pld}且{pld}→{name},但是考慮 遞移函數相依的定義:若存在X→Z和Z→Y,且Z不為超級鍵,則 稱X→Y為遞移函數相依。因為{pld}是屬於Member的身分證字號, 已是唯一的資料,也就是超級鍵。所以{mld}→{name}並不是遞 移函數相依。因此Member關聯滿足3NF

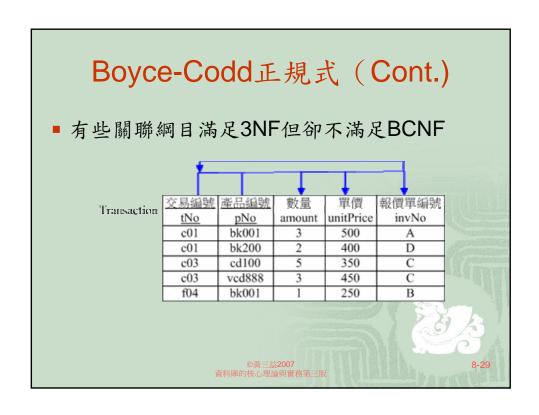
©黃三益2007 資料庫的核心理論與實務第三版 8-27

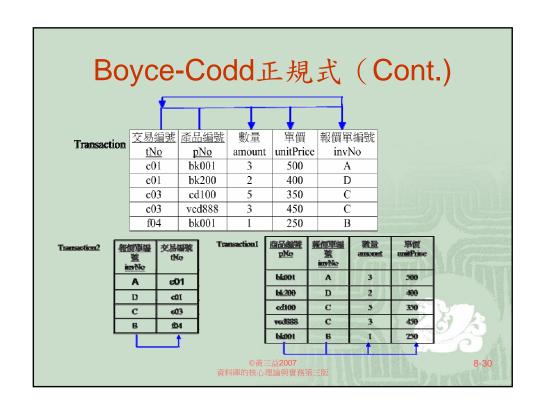
Boyce-Codd正規式

- 一個關聯綱目R滿足BCNF,如果
 - □對於R的每一個函數相依A→B,A都是超級鍵
- 滿足BCNF的關聯綱目必然也滿足2NF和3NF

 - 兩若R不滿足3NF,則存在著非鍵屬性Y是遞移函數相依於關聯鍵X。也就是有X→Z和Z→Y,且Z不是超級鍵。所以Y可由非超級鍵X、所決定

©黃三益2007 資料庫的核心理論與實務第三版





Boyce-Codd正規式 (Cont.)

- 上頁Transaction關聯滿足3NF(但不滿足BCNF)
- 上頁Transaction1和Transaction2滿足BCNF
- 何者較好?
 - 考慮所有函數相依
 - 1. $(tNo, pNo) \rightarrow amount$
 - 2. (tNo, pNo) →salePrice
 - 3. $(tNo, pNo) \rightarrow invNo$
 - 4. invNo→tNo
 - 5. $(invNo, pNo) \rightarrow amount$
 - 6. (invNo, pNo) →salePrice

©黃三益2007 資料庫的核心理論與實務第三版 8-31

Boyce-Codd正規式(Cont.)

- 只有Transaction關聯時,設定主鍵可決定
 - $(tNo, pNo) \rightarrow amount$
 - (tNo, pNo) →salePrice
 - $(tNo, pNo) \rightarrow invNo$
- 分解成Transaction1和Transaction2關聯時,設定主鍵可決定
 - $(invNo, pNo) \rightarrow amount$
 - \bowtie (invNo, pNo) \rightarrow salePrice
 - invNo→tNo
- 當分解成Transaction1和Transaction2關聯時,若仍想維持函數相依(tNo, pNo)→amount ,則需做跨關聯的檢查
- **保留函數相依的分解**:關聯網目分解後各函數相依仍可藉由檢查 單一關聯來確定
- BCNF常無法保留所有函數相依

©黃三益2007 資料庫的核心理論與實務第三版



有人可能認為8-12(b)的關聯網目沒有比較省空間,請修改圖8-12範例關聯裡的記錄,使得空間的節省會較多

Ans:

3NF

	`	SINI		
交易編號 tNo	產品編號 pNo	數量 amount	單價 unitPrice	報價單編號 invNo
c01	bk001	3	500	A
c01	bk200	2	400	A
c01	cd100	5	350	A
c01	vcd888	3	450	A
c01	bk005	1	250	A

BCNF

報價單編號 invNo	產品編號 pNo	數量 amount	單價 unitPrice
A	bk001	3	500
A	bk200	2	400
A	cd100	5	350
A	vcd888	3	450
A	bk005	1	250

 報價單編號 invNo
 交易編號 tNo

 A
 c01

©黃三益2007 資料庫的核心理論與實務第三版 8-33

第四正規式

- 第四正規式 (簡稱4NF) 是源於多值相依的概念
 - 図 多值相依X→Y:給定一個X的屬性值,便有一組Y的屬性值
 - pNo*author
 - pNo»transaction

商品編號 pNo	<u>創作者</u> <u>author</u>	交易 transaction
c01	ABC	50
c01	XYZ	42
c01	ABC	42
c01	XYZ	50
m05	DEF	51
m05	XYZ	33
m05	DEF	33
m05	XYZ	51

©黃三益2007 資料庫的核心理論與實務第三版

第四正規式 (Cont.)

ProductAuthor

	•
商品編號 pNo	<u>創作者</u> author
c01	ABC
c01	XYZ
m05	DEF
m05	XYZ

ProductTransaction

1	商品編號 pNo	交易 transactio <u>n</u>	
	c01	50	
	c01	42	
	m05	51	
	m05	33	

- 在ProductAuthor關聯裡的多值相依pNo*author和 ProductTransaction關聯裡的多值相依 pNo*transaction稱為微不足道的多值相依
- 一個關聯綱目R滿足4NF,如果 □ 每一個R的非微不足道的多值相依其左方都是超級鍵
- 滿足4NF也必定滿足BCNF

©黃三益2007 資料庫的核心理論與實務第三版 8-35

第五正規式

- 一個關聯綱目R滿足5 NF:
 - □ R無法再被分解成數個關聯R1, R2, ···, Rk, 使得R1*R2* ···*Rk=R, 或
 - □ R可以被分解成數個關聯R1, R2, ···, Rk, 使得 R1*R2* ···*Rk=R, 但是每一個Ri, 1≤i≤k, 都是 R的超級鍵。
- 5NF限制條件的發現和檢查不易,也因此在 實作上通常不予考慮

©黃三益2007 資料庫的核心理論與實務第三版