

# 《資料庫應用》

一、有一個與租船有關的資料庫，其綱要 (Schema) 及案例內容 (Instance) 如下：

Sailors(sid:integer, sname:string, rating:integer, age:real)

Boats(bid:integer, bname:string, color:string)

Reserves(sid:integer, bid:integer, day:date)

sid	sname	rating	age
22	Wawrinka	10	30
29	Brutus	6	33
31	Lubber	8	55
32	Lu	9	32
58	Rusty	8	35
64	Nishikori	10	26
71	Dustin	10	30
74	Murray	10	28
85	Nadal	10	29
95	Bob	3	60

圖一、Sailors表格的案例

sid	bid	day
22	101	6/7/2015
22	102	6/5/2015
22	103	6/3/2015
22	104	6/1/2015
32	102	5/28/2015
32	103	6/1/2015
64	104	6/3/2015
85	101	5/30/2015
85	102	6/3/2015
85	103	6/3/2015

圖二、Reserves表格的案例

bid	bname	color
101	Intelake	blue
102	Clipper	red
103	Marine	green
104	Blast	red

圖三、Boats表格的案例

請回答下列各題：(每小題5分，共40分)。請注意：答案必須具備一般性，表格案例內容只為參考用，案例內容更改後，答案仍須正確。

- (一)請劃出此資料庫相對應的實體關聯圖 (ER Diagram)。
- (二)請指出那幾個表格中的那幾個欄位，需要滿足參考限制 (Referential Integrity)？
- (三)寫出一個SQL的查詢命令，查詢“Wawrinka”所租過的船的船名。
- (四)寫出一個SQL的查詢命令，查詢rating大於或等於9的船員人數。
- (五)寫出一個SQL的查詢命令，查詢曾經租過編號為103號的船的船員名字。
- (六)寫出一個SQL的查詢命令，查詢曾經租過所有顏色為紅色 (red) 的船的船員名字 (即該船員租過所有顏色為紅色的船)。
- (七)寫出一個SQL的查詢命令，查詢同一天之中租超過一艘以上 (不包括一艘) 的船的船員名字。
- (八)寫出一個關聯代數 (Relational Algebra) 運算式，以關聯代數運算子 (selection:  $\sigma$ 、projection:  $\Pi$ 、join:  $\bowtie$ 、divide:  $\div$ 、union:  $\cup$ 、intersection:  $\cap$ 、cartesian product:  $\times$ 、difference:  $-$ ) 表示，查詢租過船名為“Blast”的船的船員名字。

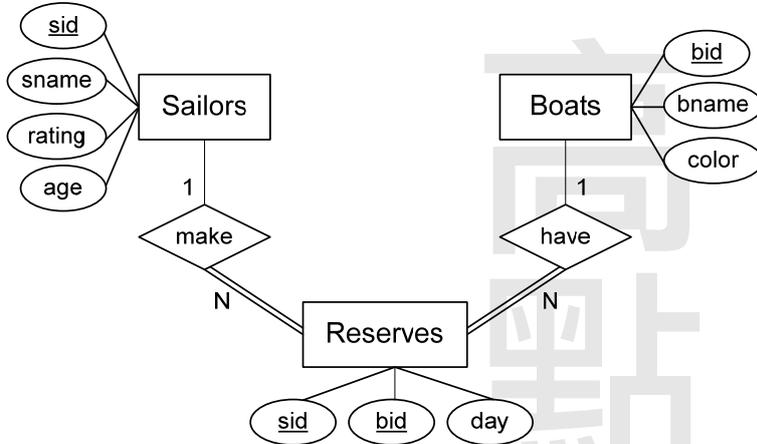
## 試題評析

今年的試題，題型不艱深，但題目較多，考生們需要把握時間，以免不及作答。用打字解題都花了不少時間，考生們用寫的，只能說辛苦了。  
第一題的ER Model, SQL幾乎是年年必考題，考生們應有所準備。關聯式代數考題出現較不頻繁，但偶爾還是會考語法，基本語法應該要熟悉。

**考點命中**

1. 《高上資料庫應用講義》第一回，唐箏編撰，第一章第七節。
2. 《高上資料庫應用講義》第二回，唐箏編撰，第五章第一節。
3. 《高上資料庫應用講義》第二回，唐箏編撰，第六章。

**答：**  
(一)



(二) Reserves的sid與bid需要滿足參考限制，即Reserve的sid與bid分別必須存在於Sailors的sid與Boats的bid。

(三) SELECT bname

```
FROM Boats B, Sailors S, Reserves R
WHERE B.bid=R.bid AND S.sid=R.sid AND sname='Wawrinka'
```

(四) SELECT COUNT(\*)

```
FROM Sailors
WHERE rating>=9
```

(五) SELECT sname

```
FROM Sailors S, Reserves R
WHERE S.sid=R.sid AND bid=103
```

(六) SELECT sname

```
FROM Sailors S
WHERE NOT EXISTS
(SELECT *
FROM Boats B
WHERE color='red' AND NOT EXISTS
(SELECT *
FROM Reserves R
WHERE R.sid=S.sid AND R.bid=B.bid)
```

(七) SELECT sname

```
FROM Sailors S, (SELECT sid, day FROM Reserves GROUP BY sid, day HAVING COUNT(*)>1) T
WHERE S.sid=T.sid
```

(八)  $\Pi_{sname}(\sigma_{bname='Blast'}(Sailor \bowtie_{Sailor.sid=Reserves.sid} Reserves \bowtie_{Reserves.bid=Boats.bid} Boats))$

二、請回答下列和資料倉儲相關的問題：

(一)請舉一個資料倉儲中的星狀資料綱要 (Star Schema) 的例子。(10分)

(二)請問資料倉儲架構中的ETL操作指的是什麼?(5分)

**試題評析**

資料倉儲基本特性考生們應該都很熟悉。星狀綱要、雪花綱要、ETL，考試出現次數較少，不過有上總複習課程的同學有加強印象，應較能作答。

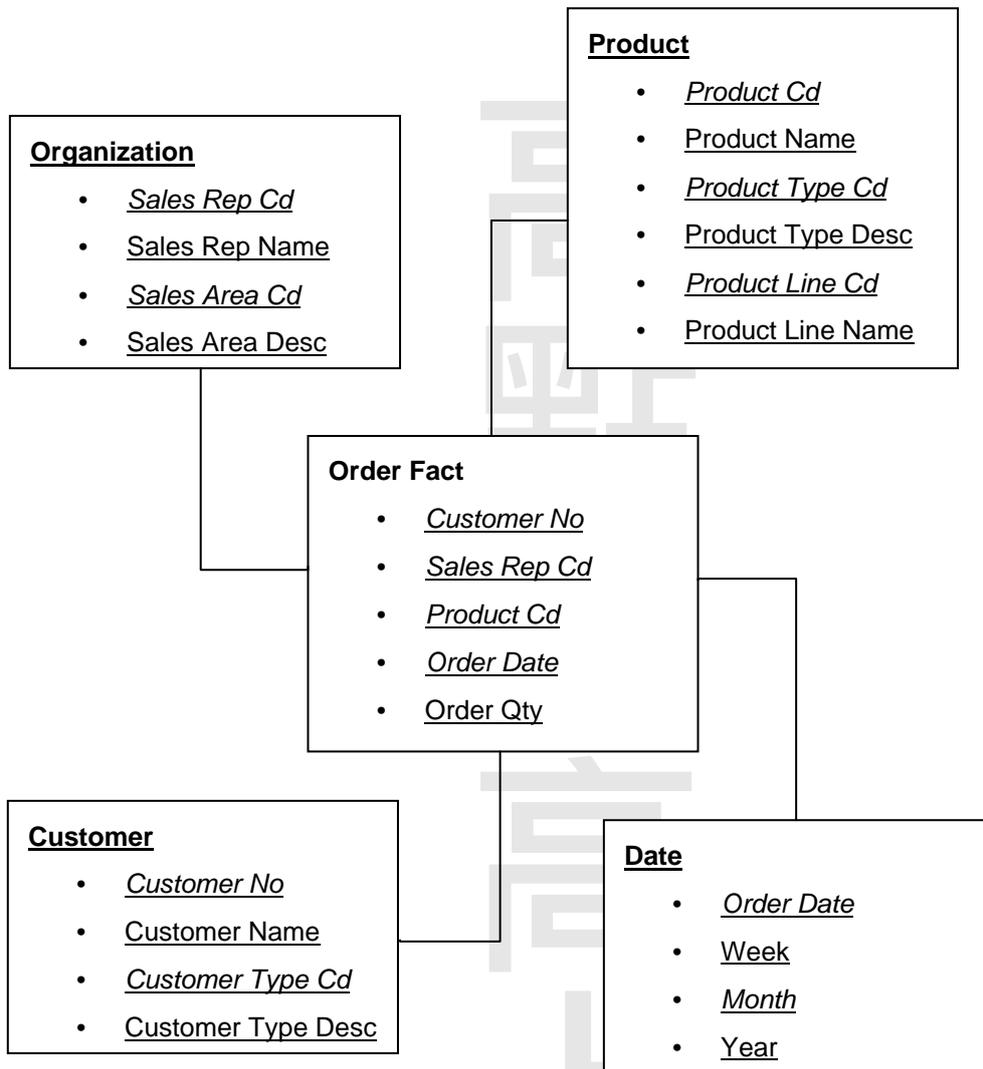
【版權所有，重製必究！】

## 考點命中

- 1.《高上資料庫應用總複習講義》，唐箏編撰，頁10-13。
- 2.《高上資料庫應用講義》第三回，唐箏編撰，頁79及上課筆記。

答：

(一)



(二)ETL為Extract Transform Load的縮寫。ETL操作即將資料由來源資料庫或檔案萃取不同資料型式定義的資料、將其彙整定義後，載入資料倉儲。ETL操作分別為：

- 1.Extract資料擷取：擷取不同平台或異質資料庫的資料檔案。
- 2.Transform資料轉換：整理所擷取出的資料，並重新定義資料格式以便載入資料倉儲。例如：資料的清理、驗證、整合、格式轉換、反正規化、彙總等。
3. Load資料載入：把轉換後的資料載入至資料倉儲，後續並定期重新載入與更新。

三、在高階語言（如C語言）中要查詢或異動資料庫時，通常會使用到一種被稱為Cursor的結構，請說明Cursor的功用及其相關的操作。（10分）

## 試題評析

Cursor很多年沒考了，過去曾經考過幾次解釋名詞，今年放成單獨一題來考，不是太有價值的題目。

【版權所有，重製必究！】

**考點命中** 《高上資料處理講義》第三回，唐箏編撰，頁101。

**答：**

在資料庫中，cursor為用來控制與連續處理資料庫回傳結果集合(result set)的一種結構。cursor可以處理資料庫系統查詢結果的單一個別紀錄，因為許多程式語言通常為程序性的而不具有一次處理整個結果集合的機制；因此結果集合的每一列必須被程式依序處理，故可以使用cursor來作指向與移動結果集合個別列的工具，將查詢結果一次傳回一筆紀錄。Cursor一般都會有open, fetch next, close等操作。

四、表格的正規化 (Normalization) 是透過切割表格 (Decomposition) 來進行，請回答下列各題：

- (一) 定義何謂無損切割 (Lossless Decomposition) ? (5分)  
 (二) 若有一個表格綱要 $r(A,B,C,D,E)$ 的函數相依關係集合 $F$ 為：

$$\begin{aligned} A &\rightarrow BC \\ CD &\rightarrow E \\ B &\rightarrow D \\ E &\rightarrow A \end{aligned}$$

請問將 $r$ 切割為 $r_1(A,B,C)$ 及 $r_2(A,D,E)$ 的切割是否是無損切割？若是無損切割，則請證明之。(10分)

**試題評析** 無損失切割，其實就是無損失合併。第二小題在課堂上有做過幾次練習，有複習就有印象。

**考點命中** 《高上資料庫應用講義》第二回，唐箏編撰，頁19、44、45及上課筆記。

**答：**

(一) 若關聯表 $R$ 被「分解(decomposition)」成數個關聯表 $R_1, R_2, \dots, R_n$ ，可以再透過「合併(join)」 $R_1, R_2, \dots, R_n$ 得到相同的資訊 $R$ 。則表示 $R$ 有一個可以分解為 $R_1, R_2, \dots, R_n$ 的無損切割。

(二) 1. 建立一matrix  $S$ ，每一關聯為一列。有該屬性者標示 $a$ ，表示該關聯候選鍵可決定之屬性；其餘欄位標示 $b$ 。

Matrix  $S$ ：

	A	B	C	D	E
r1	a	a	a	b	b
r2	a	b	b	a	a

2. 對於每個功能相依 $X \rightarrow Y$ ，若任兩列決定因素 $X$ 值相同(皆為 $a$ )，則可設定其 $Y$ 值亦同。表示經由該功能相依可額外決定之 $Y$ 屬性。

	A	B	C	D	E
r1	a	a	a	b	b
r2	a	<del>b</del> a	<del>b</del> a	a	a (因 $A \rightarrow BC$ )

3. 由上， $r_2$ 整列皆為 $a$ ，表示 $r_1$ 與 $r_2$ 可透過功能相依 $A \rightarrow BC$ 無損失合併為原關聯。  
故 $r_1$ 與 $r_2$ 為無損切割。

五、線上交易處理 (Online Transaction Processing) 是資料庫系統中的一個重要功能，請回答下列各題：(每小題5分，共20分)

- (一) 定義什麼是一個交易 (Transaction) ? 它必須滿足那四個特性?  
 (二) 何謂並行控制 (Concurrency Control) ? 資料庫中若無並行控制的機制，則可能會產生什麼問題?  
 (三) 何謂兩階段鎖定 (Two-Phase Locking) ? 其目的為何?  
 (四) 圖四為兩個交易 $T_1$ 及 $T_2$ 的執行過程 (Schedule)，假設交易執行的過程中有實施兩階段鎖定，請問這個執行過程 (Schedule) 的結果是順利執行完畢，或是發生死結狀態 (Deadlock) ?

【版權所有，重製必究！】

T1	T2
Read(A)	
A:=A-50	
Write(A)	
	Read(B)
	B:=B-10
	Write(B)
	Read(A)
	B:=A+10
	Write(A)
Read(B)	
B:=B+50	
Write(B)	

圖四、交易T1、T2的執行過程（指令的上下位置表示執行時間的先後，上面的指令比下面的指令先發生）

<b>試題評析</b>	交易管理考題，每題都先來個解釋名詞，需要花時間花墨水作答，但屬於送分題型。最後一題的2PL也算基本，好好作答不難取分。
<b>考點命中</b>	《高上資料庫應用講義》第三回，唐箏編撰，頁1-2、17-18。

**答：**

- (一)交易指一個存取或改變資料庫內容的執行，為一件工作的邏輯單位。交易必須滿足ACID四特性，即基元性(Atomicity)、一致性(Consistency)、孤立性(Isolation)、永久性(Durability)
- (二)讓多個交易可以在同一時間存取同一筆資料，稱之為並行；並行控制即讓交易在並行的狀況下運作，而不會互相干擾，確保交易間的孤立性(Isolation)，以及提高交易的效率。資料庫若無並行控制機制，可能產生如遺失更新(Lost Update)、不一致分析(Inconsistent analysis / Incorrect summary)、或未委任相依(Uncommitted dependency / Dirty reads / Temporary update)等問題。造成資料庫的結果不正確。
- (三)兩階段鎖定，簡稱2PL，限定交易中所有鎖定動作(含升級)，包括Read-lock與Write-lock，必須在所有解除鎖定(含降級)Unlock動作之前。鎖定(含升級)動作稱作擴展階段，解除鎖定(含降級)則稱為收縮階段。2PL的目的即為交易的並行控制，並可確保排程的可序列化與正確性。
- (四)該過程會產生死結狀態。因交易T1在擴展階段進行了Read-Lock(A)、Write-Lock(A)但無法取得Read-Lock(B)，因B已被T2鎖定；反之T2在擴展階段進行了Read-Lock(B)、Write-Lock(B)但無法取得Write-Lock(A)，因A已被T1鎖定。加入鎖定的過程操作順序如下：

T1	T2
Read-Lock(A)	
Read(A)	
A:=A-50	
Write-Lock(A)	
Write(A)	
	Read-Lock(B)
	Read(B)
	Write-Lock(B)
	Write(B)
	Read-Lock(A)→操作不允許
	.....
Read-Lock(B) →操作不允許	

兩排程皆仍在擴展階段，【尚未進入收縮階段，故無法解除鎖定。但彼此互相等待！】，造成死結狀態。