

《資料庫應用》

一、假設某一機關的關聯資料庫中有三個表 (Table)，分別是職員資料表 (E)、處室單位表 (D) 以及請假紀錄表 (H)，其中，依序各表之欄位有：

表 E (職員編號，職員姓名，單位代號，工作年資)

表 D (單位代號，單位名稱，單位主管編號，單位任務)

表 H (職員編號，請假別，請假開始日期，請假結束日期，職務代理人編號)

其中，單位主管也是屬於職員的一類，單位主管編號就是其職員編號；每位單位主管或職員只能歸屬於某一單位；職員 (含單位主管) 請假時，必須由同一單位的其他職員代理其業務，且一位職員至多同時可代理兩位同單位請假職員 (含單位主管) 的業務，單位主管不得代理請假職員之業務：

(一) 為確保上述資料庫中資料的正確性與一致性，請問應該分別建置那些主鍵 (primary key) 與外部鍵 (foreign key)？(5 分)

(二) 請寫出 SQL 語句，可從上述資料庫中，找出職員編號：105，在整個去年 (西元 2015 年) 所有的請假中，是由職員編號：203 所代理的請假紀錄資料，亦即：(請假別，請假開始日期，請假結束日期)。(10 分)

(三) 請寫出 SQL 語句，當職員編號：105，擬於 2016/9/8 至 2016/9/9 請休假，從資料庫中查詢可代理其業務同仁的職員編號。(10 分)

試題評析	前兩題為基本常見考題，應完整取分。第三小題語法稍為複雜，需花些心思。
考點命中	《高點·高上資料庫應用講義》第二回，唐箏編撰，頁 98-109。

【擬答】

(一)

職員表(E)：

主鍵 - 職員編號

外部鍵 - 單位代號

處室單位表(D)：

主鍵 - 單位代號

外部鍵 - 單位主管編號

請假記錄表(H)：

主鍵 - (職員編號，請假開始日期)

外部鍵 - 職員編號

職務代理人編號

(二)

SELECT 請假別，請假開始日期，請假結束日期

FROM H

WHERE 職員編號=105 AND 職務代理人=203

(三)

SELECT 職員編號

FROM E

WHERE 單位代號=

(SELECT 單位代號

FROM E

WHERE 職員編號=105)

AND 職員代號 NOT IN

【版權所有，重製必究！】

```
(SELECT 職務代理人編號
FROM H
WHERE 請假開始日期<=2016/9/9 AND 請假結束日期>=2016/9/8
GROUP BY 職務代理人編號
HAVING COUNT(*)>=2)
AND 職員代號 NOT IN
(SELECT 單位主管編號
FROM D)
```

二、假設一關聯網目 (Relational Schema) $R(S, E, G)$ 中有 3 個屬性，其中， $\{S, E\}, \{S, G\}$ 是候選鍵 (Candidate Keys)，功能相依 (Functional Dependency) 有：

$\{SE \rightarrow G, G \rightarrow E\}$

(一) 請問 R 滿足之最高正規型式 (Normal Form) 為何？並請說明原因。(15 分)

(二) 請問如何處理 R ，使它能再滿足更高一層的正規型式。(10 分)

試題評析	正規化為常見題型，題目不複雜，應不難取分。
考點命中	《高點·高上資料庫應用講義》第二回，唐箏編撰，頁 14-16。

【擬答】

(一) R 滿足之最高正規型式為 3NF。

因關聯 R 無多值屬性、複合屬性，為 1NF；

取 $\{S, E\}$ 為主鍵，無部分功能相依於主鍵之屬性，為 2NF；

不存在任何遞移相依，為 3NF；

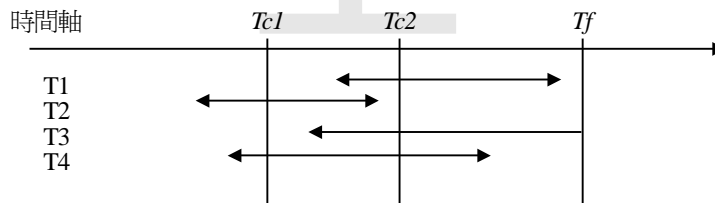
但決定因素 G 不為關聯之候選鍵，故非 BCNF。

(二) 可將關聯分解為如下 R_1 及 R_2 ，即皆滿足 BCNF。

$R_1(S, G)$

$R_2(E, G)$

三、假設有 4 筆同步執行的交易 (transactions)，其執行狀態如下圖，其中，「←」代表交易開始執行，「-」代表該交易執行中，「→」代表該交易正常結束執行； $Tc1$ 與 $Tc2$ 表示資料庫系統執行檢查的兩個時間點 (checkpoint)， Tf 則是表示在該時間點時，資料庫系統發生故障 (system failure)。請問要用何種機制，來達成資料庫系統的回復 (Recovery)？請詳細說明此機制如何運作，以及分別在各時間點上，應該進行那些事項與訊息紀錄，並說明每一筆交易在發生系統故障 (Tf) 後，該進行何種動作？(25 分)



試題評析	交易管理的復原機制。題目未說明要用哪種復原機制，因此就都列出吧。答題時必須說明採用哪種演算法。
考點命中	《高點·高上資料庫應用講義》第三回，唐箏編撰，頁 46-52。

【擬答】

系統故障 Tf 時，交易

T_3 尚未 commit

T_1 與 T_4 已 commit 但尚未經過任何 checkpoint

T_2 已 commit 且已經過 checkpoint Tc_2

若該排程採用

- (一)延遲更新技術，即 No-UNDO/REDO 演算法，則尚未 commit 的 T3 資料不會寫入資料庫，可忽略不處理；已 commit 但尚未經 checkpoint 的 T1 與 T4 的所有寫入操作，必須重新執行(REDO)；已 commit 且經 checkpoint 的 T2 資料已進資料庫可忽略不處理。
- (二)立即更新技術，即 UNDO/REDO 或 UNDO/NO-REDO 演算法(以下以 UNDO/REDO 為例)，則尚未 commit 的 T3 資料可能已寫入資料庫，必須 UNDO 復原所有操作；已 commit 但尚未經 checkpoint 的 T1 與 T4 的所有寫入操作，必須重新執行(REDO)；已 commit 且經 checkpoint 的 T2 資料已進資料庫可忽略不處理。

四、請說明何謂大數據 (Big Data) ? (5 分) 何謂資料探勘 (Data Mining) ? (5 分) 近年來，各級政府與部門機關推動政府資料開放 (Open Data)，在推動的過程中，主辦機關應該注意那些重要原則? (15 分)

試題評析	此題為近年來政府重點發展方向。考生除教科書外，可多加留意政府推行的資訊科技相關政策，不時會出現與政府發展方向相近之考題。
考點命中	1.《高點·高上資料庫應用講義》第二回，唐箏編撰，頁 85。 2.《高點·高上資料庫應用講義》總複習，唐箏編撰，頁 14。

(一)Big Data：中文常稱作海量資料、巨量資料、大數據等。Big Data 是大量(Volume)、高速(Velocity)、多變(Variety)的資訊資產，構成了複雜(Complexity)的結構與挑戰。它需要新型的處理方式去促成更強的決策能力、洞察力、與最佳化處理。Big Data 的 3V 特性：(1) Volume (容量)：大量的數據；(2) Velocity (速度)：快速地處理數據；(3) Variety (多樣)：內容豐富的數據。

(二)Data Mining：簡單來說，資料探勘(Data Mining)就是從大量資料中提取(extract)或稱探勘(mining)出知識的一種方法。Data mining 以自動或半自動的方式對大量資料作分析，以找出有意義的關係或規則。

(三)政府 Open Data 的八大原則：

- 1.全面開放 (Complete)：所有公眾資料都可被取得。所謂公眾資料，是指 (與眾人相關但) 不受隱私、安全、或其他特別權利合理限制的資料。
- 2.資料完整 (Primary)：資料以其來源最完整型式、最精細的尺度收錄，不使用加總或修改後的格式。
- 3.即時發布 (Timely)：資料應就需要儘速提供，以保留資料價值。
- 4.便民使用 (Accessible)：資料應向最廣泛的使用者揭示，並容許最大範圍的使用。
- 5.機器處理 (Machine processable)：資料須組織為可被自動化處理的形式。
- 6.不作限制 (Non-discriminatory)：資料公開給任何人使用，不需要註冊。
- 7.公開格式 (Non-proprietary)：用以發布資料的格式不可被特定人所控制。
- 8.不需授權 (License-free)：資料不受著作權、專利權、商標權，以及營業秘密的管制；然在隱私、安全，或其他特別權利上可受合理限制。

【版權所有，重製必究！】