

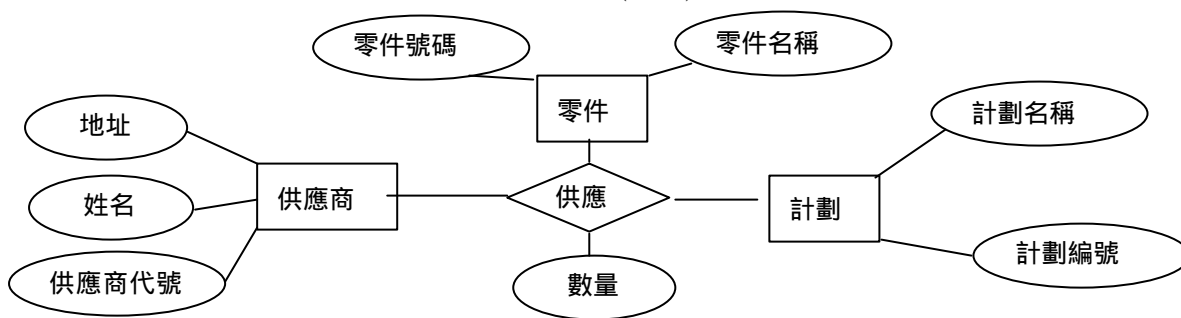
# 《資料處理》

## 試題評析

本次資處的考題主要涵蓋範圍為資料庫及電子商務兩個領域，與今年普考試題範圍幾近相同，對照近三年來的高考試題差異不大，主要差異處只是物件導向技術。第一題的實體關係圖及第二題的完整性可屬一般試題，第三題的正規化試題及第四題的SSL及SET的差異則是一試常見的題型，只有第五題的Data Mining較不易作答；本份試題難度適中，預估中間考生應有60分，實力佳者應可有80分。

依據近幾年試題範圍及目前資訊網路運用情形來看，未來本科考試趨勢應仍以資料庫、資訊網路(含Web Services及電子商務)、物件導向技術(含物件導向資料庫)為主。

一、請將下列之實體關係圖利用關係綱目來表示。(20分)



答：

共得四個關聯表如下：

1. 供應商關聯表(供應商代號, 姓名, 地址), 主鍵為(供應商代號)
2. 零件關聯表(零件號碼, 零件名稱), 主鍵為(零件號碼)
3. 計劃關聯表(計劃編號, 計劃名稱), 主鍵為(計劃編號)
4. 供應關聯表(供應商代號, 零件號碼, 計劃編號, 數量), 主鍵為(供應商代號, 零件號碼, 計劃編號)三者複合而成

二、資料庫的瞬間值必須滿足那四種完整性約束？試解釋說明之。(20分)

答：

四種完整性約束如下：

(一)型態(type)約束：指定屬性的合法值。基本上，型態約束只是(或邏輯上相當於)該型態合法值的列舉而已。例如：一個簡單的例子，WEIGHT型態的型態限制為：

```
TYPE WEIGHT POSSREP (RATIONAL)
```

```
CONSTRAINT THE_WEIGHT (WEIGHT) > 0.0;
```

這個例子限制了重量(Weight)為可以由有理數(rational number)來表示而且大於0。如果運算式的計值結果不滿足該限制，則該運算式將失敗。

(二)屬性(attribute)約束：指定屬性的合法值。基本上，屬性約束只是將一個屬性宣告成某一種型態而已。例如：考慮供應商關係變數定義：

```
VAR S BASE RELATION
```

```
{ S#      S#,
  SNAME  NAME,
  STATUS INTEGER,
  CITY   CHAR } . . . ;
```

在這個關係變數中，屬性S#、SNAME、STATUS和CITY的值分別被限制成型態S#、NAME、INTEGER和CHAR。

(三)關係變數(relVar)約束：指定關係變數的合法值。關係變數約束是針對個別關係變數的限制(它是由關係變數表示而成，不過可能具有任意複雜度)，例如：

```
CONSTRAINT PC4
```

```
IS_EMPTY ( P WHERE COLOR = COLOR ( 'Red' ) )
```

```
AND CITY 'London' );
```

(“紅色零件必須存放在倫敦”)。

關係變數限制一定是立刻檢查的(因為任何敘述的任何部份的執行結果都可能違反限制)，因此，如果有任何敘述企圖指定一個值給某一關係變數而使得它違反其關係變數限制，則該敘述將被拒絕。

(四)資料庫(database)約束：指定屬性的合法值。資料庫約束是和多個不同關係變數有關的限制，例如：

```
CONSTRAINT DBC1
IS_EMPTY ( ( S JOIN SP
WHERE STATUS < 20 AND QYT > QTY (500) );
```

(“狀態值小於20的供應商所供應的任何零件數量都不可以超過500”)

三、 $R=\{S, A, I, P, W\}$ ， $F=\{S \rightarrow A, SI \rightarrow P, I \rightarrow W\}$ 試找出一個Lossless Join Decomposition into BCNF Relation Schemas。(20分)

**答：**  
BCNF的定義為：若且唯若，一個關聯表的所有Determinant都是該關聯表的Candidate key(依選鍵)。

將本題的關聯表R進行Lossless Join Decomposition 為BCNF後的結果，成為三個關聯表，分別為R1，R2，R3，說明如下：

$R_1 = \{S, A\}$ ， $F_1 = \{S \rightarrow A\}$ ，Determinant=S，Candidate Key=S  
 $R_2 = \{I, W\}$ ， $F_1 = \{I \rightarrow W\}$ ，Determinant=I，Candidate Key=W  
 $R_3 = \{S, I, P\}$ ， $F_1 = \{SI \rightarrow P\}$ ，Determinant=(S,I)，Candidate Key=(S,I)

因上述的R1, R2, R3的Determinant都是Candidate Key，故R1, R2, R3均符合BCNF。

四、請比較SSL與SET之交易區別。(20分)

**答：**  
(一)SSL(Secure Socket Layer)

- 用途：是一種網際網路上的安全協定，也是電子商務網站上，最常使用的一種資料傳輸安全形式。採用SSL的主要目的是要讓使用者在網路上所傳輸的每一份資料，都能受到安全的保護，包括：資料的加密功能、身份認證及資料的檢核。
- 運作方式：當SSL啟動時，客戶端端的瀏覽器會傳送本身的公開金鑰(public Key)給網站伺服器，而網站伺服器端則會將一個私密金鑰(Private Key)傳送到客戶端的瀏覽器上，雙方建立起一個安全資料交換的傳輸環境。
- 推展上的問題：包括缺乏認證機構、無法確認消費者是否為持卡本人以及店家對客戶資料的非法使用等問題。

(二)SET( Secure Electronic Transaction，安全電子交換協定)

- 用途：是由Visa與Master Card兩大信用卡組織提出的安全電子交易協定。此協定是一種應用在網際網路上，以信用卡為基礎的電子付款系統規範。SET規格是使用公開金鑰(Public key)所編成的密碼文件，以確保開放性網路上持卡交易的安全性。
- 運作方式：為使Internet上的交易兼具安全與方便性，必須規範所有在Internet進行交易個體的作業方式及流程。SET協定的系統架構中存在四種個體角色，包括：持卡人、電子特約商店、收單銀行、認證中心及發卡銀行，以確保電子交易進行的安全防護。
  - 持卡人先向發卡銀行申請信用卡，同時向相關認證機構取得數位簽證。
  - 持卡人使用瀏覽器連結至電子網站上購物。
  - 結帳時，該網路商店將通知持卡人端的電子錢包軟體，要求持卡人填入其信用卡的相關資料(例如：信用卡號碼或有效期限等)。
  - 電子錢包軟體將這些資料加密，並透過網際網路傳送至販售商品的商店軟體上。
  - 商店再將所收到的資料，整合成授權付款訊息，傳送至信用卡取款銀行。
  - 取款銀行再循原有的信用卡清算網路系統，進行貨款的支付作業與持卡人的帳單寄發。
- 推展上的問題：處理速度較慢。

五、請分別說明下列Data Mining可以建立之六種模式：Classification、Regression、Time-Series Forecasting、Clustering、Association、Sequence Discovery。(20分)

**答：**  
Data Mining可以建立的六種模式說明如下：

- (一)Classification：指由交易資料的屬性來做目標的分類。例如：汽車經銷商可能會將顧客記錄分成高級車、中級車及普通車三種目標。Classification(分類)系統可能會產生一套規則，認為職業屬性為商業，收入屬性為超過10萬元，年齡介於45至55之間，他應具有高級車身價。
- (二)Regression：迴歸是利用一對相關的變數來預測其中第二個變數值的過程。最普遍使用的迴歸模式是線性迴歸。例如可依據顧客歷次的消費金額及社會經濟發展現況，預估其未來的消費能力。
- (三)Time-Series Forecasting：指的是商品的特定時間順序，通常歷時性分析首先要能知道顧客的個人資料，然後我們才能夠將不同時間的各筆交易結合起來。
- (四)Clustering：係根據某種特性，將資料庫分成多個子集合或Cluster(聚集)。例如：根據顧客過去消費習慣、教育程度、年

收入等特性，找出潛在顧客。

- (五)Association：考量顧客於同一次交易所購買的多項產品間是否有特定購買模式存在。這些模式可能由於顧客消費習性，或由於產品間的關連性甚或其它一些外在誘因(例如：某些特定商品大特價)等因素而造成。經由關聯規則資料萃取，可以發現同一交易時點，某一類的產品與另一類的產品會一起被購買。例如：買鐵鎚的顧客通常會買鐵釘。
- (六)Sequence Discovery：係指根據顧客過去數次的購物行為，做因果分析。例如：顧客購買燃油燈後，下一次他會購買燃油添加物，這二項產品具有因果關係。