

# 《資料庫應用》

一、在關聯式資料庫 (relational database) 軟體所提供的功能中，有些定義可以讓系統協助檢查資料的一致性 (consistency)，有些定義則可以提升查詢處理的效率。假設要在資料庫系統中建立兩個表格，其中：學生表格用以記錄每個學生的基本資料，包含學號 (ID)、姓名 (name) 和主修系別 (dname)；系所表格用以記錄每個系的基本資料，包含系名 (dname)、系館名稱 (building) 和系主任 (chair)。此二表格的定義如下所示：

student (ID, name, dname)

department (dname, building, chair)

- (一) 請問你贊同把屬性 ID 定義為 student 表格的主鍵 (primary key) 嗎？請詳述你贊同或不贊同的原因。(10 分)
- (二) 假設屬性 dname 已經被定義為 department 表格的主鍵，你贊同把 student 表格中的屬性 dname 定義為外來鍵 (foreign key) 參照到 department 表格嗎？請詳述你贊同或不贊同的原因。(10 分)
- (三) 你贊同把屬性 ID 定義為索引 (index) 嗎？請說明你贊同或不贊同的原因。(10 分)

試題評析	本題考主鍵、外來鍵、索引，題目並不難，可說是基本分，了解主鍵、外來鍵、索引特性的同學應該能順利作答。
考點命中	《高點·高上資料庫講義》第一回，唐箏編撰，頁44-45。

## 【擬答】

(一) 贊同

主鍵必須具備有 1. 唯一性 2. 最小性 3. 不可為 null

在 student 表格中屬性 name & dname 不具有唯一性，只有 ID 能符合要求。

(二) 贊同

因為 dname 在表格 department 為主鍵，所以 student 表格的 dname 可以參照到 department 表格，但要注意所有 student 表格的 dname 都要在 department 表格有對應值。

(三) 贊同

索引能加快查找主鍵的速度，一般資料庫都會這樣建索引。不過如果資料庫資料會常需要大量的新增刪除，索引得常跟著修正，會比較不適合。

二、假設資料庫系統中定義了老師的授課資料，其中 instructor 表格定義了老師的教職員編號 (IID)、姓名 (name) 和系別 (dname) 等三個屬性，而 teach 表格定義了老師的教職員編號 (IID)、所教授的課程代碼 (CID) 和授課學年度。兩個表格如下所列，注意所有屬性皆為字串型態，且底線標示了各表格的主鍵 (primary key)：

instructor (IID, name, dname)

teach (IID, CID, year)

- (一) 請寫出 SQL 查詢句 (query statement)，列出每個系的系名以及該系的授課老師個數。(10 分)
- (二) 請寫出 SQL 查詢句，列出名為「張三」的老師，於 106 學年度所有開設的課程代碼。(10 分)
- (三) 請問以下 SQL 查詢句是為了找出什麼樣的資料？請用中文詳述其資料語意 (semantics)：(10 分)

SELECT IID, CID

FROM teach

WHERE year = '107' and IID in(SELECT IID FROM instructor WHERE dname = '資工');

<b>試題評析</b>	本題考基本的 SQL 語法，也是基本題，同學應該都能順利作答。
<b>考點命中</b>	《高點·高上資料庫講義》第二回，唐箏編撰，頁 99-130。

**【擬答】**

(一)

```
SELECT dname, COUNT(*) AS 授課老師個數
FROM instructor
GROUP BY dname
```

(二)

```
SELECT T.CID
FROM instructor I, teach T
WHERE I.IID=T.IID AND T.year='106' AND I.name='張三'
```

(三)此 SQL 語法是為了從 teach 表格中輸出資料中的教職員編號(IID) 與所教授的課程代碼(CID)兩個欄位。這些資料必須符合以下兩個要求

1.year='107' -->代表 107 學年度

2.IID in ... -->教師要為資工系教師

簡單的說，就是輸出 107 學年度有授課的資工系教師的教職員編號跟課程代碼。

三、假設我們要建立一個資料庫系統，以管理本單位內員工使用交通工具的狀況。系統需求與限制如下所列：

1.每台交通工具都要記錄其牌照號碼和購買年份。

2.每個員工都要記錄其員工編號和職稱。

3.每次某個員工使用某台交通工具後，都要記錄其使用的日期。

4.員工只要有需求皆可以使用單位內的交通工具，並沒有限定使用的次數。

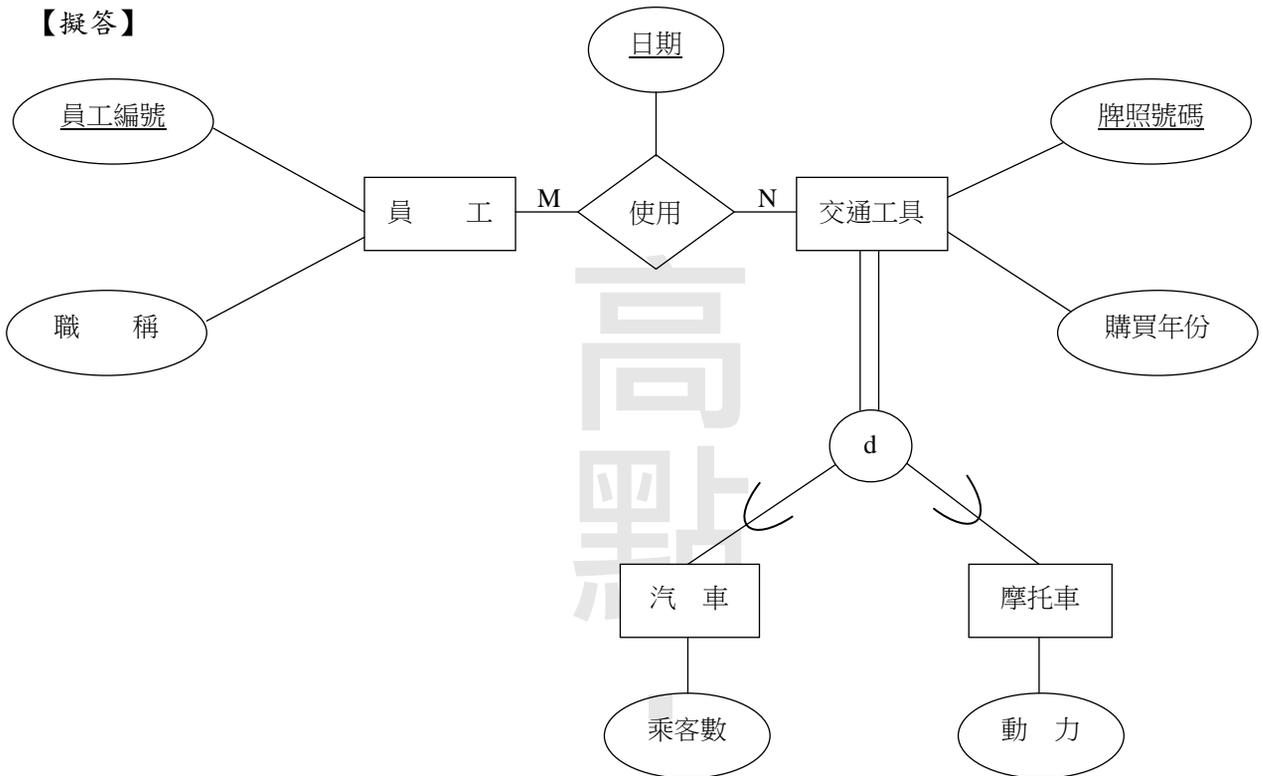
5.交通工具分為汽車和摩托車兩種，汽車要記錄其可承載之乘客數，摩托車要記錄其是否為電動車。

請根據以上描述，利用實體關係模式 (Entity-Relationship model) 進行設計，並畫出對應的實體關係圖 (ER-diagram)。請注意圖中必須標明實體集合、關係集合、屬性、主鍵和對應數量限制等，也可使用擴充實體關係模式 (Extended ER model, EER) 的表示法。(20 分)

<b>試題評析</b>	本題考 ER-diagram，有兩個要特別注意的地方：一個是員工可以多次使用同一個交通工具，所以關係上的日期屬性要成為主鍵之一；另一個地方是交通工具要用擴充實體關係模式來表達。這兩個地方老師上課都有教，同學應該不難作答。
<b>考點命中</b>	《高點·高上資料庫講義》第一回，唐箏編撰，頁 70-83。

【版權所有，重製必究！】

## 【擬答】



四、關聯式資料庫的技術已經相當成熟，也已經被一般企業或政府機構廣為使用，但是近年來又出現大數據 (Big data) 的議題，以及 NoSQL 類的資料庫軟體。

(一) 請問你認為「大數據」所欲處理的資料，與一般傳統資料庫處理的資料，有何不同之處？請列出最明顯的 3 點特性並解釋之。(10 分)

(二) 請問你認為那類軟體比較適合處理大數據？是常用的關聯式資料庫軟體，譬如 MySQL、SQL Server 等，還是 NoSQL 類的軟體，如 MongoDB、Cassandra、HBase 等，請詳述你的理由。(10 分)

**試題評析** 巨量資料有四個特性(4V)，但本題只要寫三個，建議以 Volume、Velocity、Variety 作答。

**考點命中** 《高點·高上資料庫總複習講義》第一回，唐箏編撰，頁 7-10、16。

## 【擬答】

(一) 大數據要處理的資料，與一般資料相比，有以下三個明顯不同之處：

1. Volume：指出大數據資料規模巨大，因此首先必須對產生的大量數據資料進行處理與保存。
2. Velocity：指出資料的即時性，對於資料的處理與反應的時效必須即時快速。
3. Variety：指出使用資料型態的多樣性，Big data 資料的來源包羅萬象，有文字、影音、網頁、串流資料、行動電話、感測裝置，包含結構化、半結構化以及非結構化的資料型態。

(二) 當代典型的關聯式資料庫在一些大數據資料的應用中表現了糟糕的效能；而 NoSQL 資料儲存可以不需要固定的表格模式，也經常會避免使用 SQL 的 JOIN 操作，一般有水平可延伸性的特徵。

此外，NoSQL 不使用 SQL 作為查詢語言，且不需遵守交易管理 ACID 特性，很適合用來處理有不同資料型態的大數據資料。