

《資料庫應用》

試題評析

第一、二題為資料庫傳統題型，較容易掌握分數；第三題的資訊檢索，在一般資料庫教科書中較少提及；第四題的個體關係圖並非一般資料庫幾本主要教科書的表達形式，故三、四題較不易掌握完整分數。一般同學應可取得50分上下，程度較佳同學應可取得70分以上。

一、名詞解釋：本題請依序抄題作答，作答時勿僅是「英翻中」，並儘可能舉例說明。（每小題五分，共三十分）

- (一) Integrity Checking
- (二) Candidate Key
- (三) Cartesian Product
- (四) Dead Lock
- (五) Object
- (六) Concurrency Control

答：

- (一) 一個良好的關聯式資料庫系統，必須符合關聯的資料整合性(Data Integrity)限制，故必須對資料庫從事整合性的檢查(integrity check)。資料整合性包括檢查主鍵的個體整合性限制(Entity Integrity Constraint)，以及檢查外鍵的參考整合性限制(Referential Integrity Constraint)。例如：員工資料表的身分證字號不可重複且不可為空，員工的部門代號必須在部門資料表的部門代號中找到相對應的值。
- (二) 候選鍵(Candidate key)由一個關聯的屬性子集(attribute subset)組成，可唯一識別關聯之值組。候選鍵必須具備唯一性(Uniqueness property)及最小性(Minimality)。例如：員工(員工代號，身分證字號，姓名，電話，住址)為一員工關聯表，其中員工代號及身分證字號皆為關聯之候選鍵，皆可決定關聯中其他屬性值。若選擇員工代號為主鍵，身分證字號則為替代鍵。
- (三) 卡式積(Cartesian Product)為關聯代數的一個基本運算子，以 \times 為符號，用來合併兩個關聯。無特殊條件限制，以所有可能性合併兩關聯，故傳回所有可能組合的值組。例如：

員工資料檔

| 員工代號 | 姓名 |
|-------|-----|
| 00369 | 陳一郎 |
| 00147 | 王小二 |
| 00007 | 張山峰 |

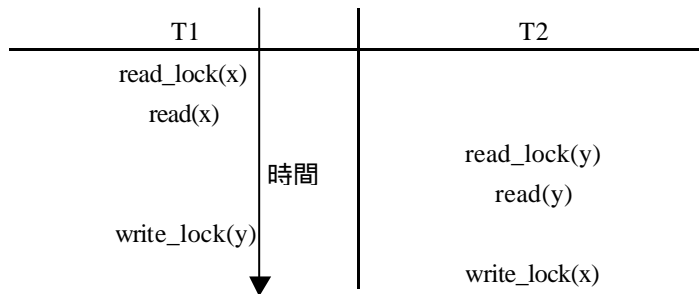
部門資料檔

| 部門代號 | 部門 |
|------|-----|
| 001 | 業務部 |
| 003 | 資訊部 |

員工資料檔 \times 部門資料檔

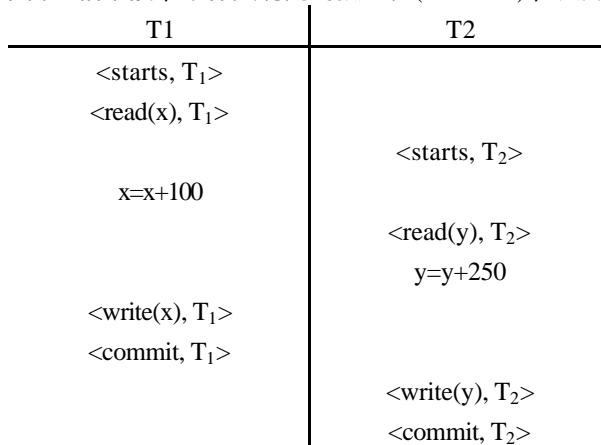
| 員工代號 | 姓名 | 部門代號 | 部門 |
|-------|-----|------|-----|
| 00369 | 陳一郎 | 001 | 業務部 |
| 00369 | 陳一郎 | 003 | 資訊部 |
| 00147 | 王小二 | 001 | 業務部 |
| 00147 | 王小二 | 003 | 資訊部 |
| 00007 | 張山峰 | 001 | 業務部 |
| 00007 | 張山峰 | 003 | 資訊部 |

- (四) 當兩個以上的交易同時進行，彼此皆佔有某些資源，但亦企圖奪取對方資源時(互相等待對方解除某項鎖定)，便會發生死結(Deadlock)的現象。形成死結的四個必要條件為：互斥現象(mutual exclusive)、握有並等待(hold and wait)、不可搶先(no pre-emption)、以及循環等待(circular waiting)。例如：下面兩交易形成死結



(五)物件(object)為一個抽象的概念，用以表示應用範圍中的各種個體；而物件之間則透過訊息(Message)來從事溝通。一個完整的物件必須包含唯一的物件識別碼(OID)、資料(屬性)、及相關的操作。例如：員工有員工代號、姓名、電話、住址等屬性，有新增員工、員工離職等操作。

(六)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(concurrency control)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。例如：下面為兩個並行之交易



二、請說明主從式架構 (Client Server Architecture) 的優點。為何還需要三層式架構 (Three Tier Architecture) ? 為何需要N層式架構 (N-Tier Architecture) 。（二十分）

答：

(一)主從式架構優點：

- 1.分散控制：可藉由client本身的運算能力，分擔server處理資料的負荷。
- 2.易於擴充：可視需求動態增加新的伺服器。
- 3.開放式架構：可支援異質性的軟硬體系統。
- 4.系統可用率增加：即使某伺服器失敗，其工作仍可由其他伺服器接手，增加系統可用率及可靠度。
- 5.資源共享：主從式架構使得資源共享變得更為容易。
- 6.友善的使用者介面：支援多媒體的圖形使用者介面(GUI)，更具使用者親合力。
- 7.降低成本：降低企業購買大型主機的投資。

(二)1.針對不同的伺服器(server)系統，需設計不同的用戶端(client)介面。

2.用戶端越多，網路連結越多，伺服器的負擔也越重。

三層式架構於主從式架構的伺服器端與用戶端之間，加上一層中間層(Middleware)，存放企業邏輯，以處理用戶端與伺服器間往來的業務。可整合後端不同的伺服器，用戶端亦只需面對單一的中間層介面，以統一的方式呈現內部的資料。

(三)三層式架構規模不斷擴大後，中間端將成為此架構的瓶頸。故提出N層式架構，將中間層作擴充，以分擔三層式架構中間層的負擔。

三、請比較資訊檢索 (Information Retrieval) 與資料探勘 (Data Mining) 的異同；請比較資料庫 (Database) 與資料倉儲 (Data Warehouse) 的異同。（二十分）

答：

(一)相同點：資訊檢索 (Information Retrieval) 與資料探勘 (Data Mining) 皆為從大量資料中找出有用資訊的技術。

相異點：

| | 資訊檢索 | 資料探勘 |
|------|--------------------------|------------------|
| 資料來源 | 多為文件式資料(document) | 資料庫(database)之資料 |
| 資料類型 | 文字(text)、多媒體(multimedia) | 格式化文字記錄、符號、數字 |

| | | |
|------|------------|-------------|
| 主要技術 | 關鍵字、相似度比對等 | 關聯、叢集、分類等 |
| 困難 | 非結構化資料的處理 | 資料探勘分析目的及方法 |

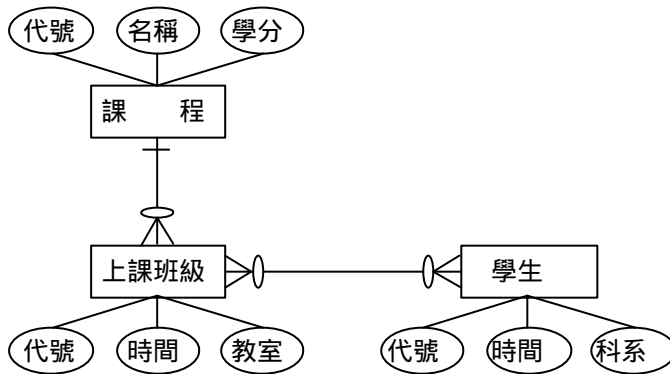
(二)相同點：資料庫 (Database) 與資料倉儲 (Data Warehouse) 皆可儲存企業中之大量資料。

相異點：

| | 資料庫 | 資料倉儲 |
|------|---------------|---------------------------------|
| 資料來源 | 個別儲存與應用的孤立資料 | 異質性的企業整合資料 |
| 時序性 | 目前(current)資料 | 目前(current)資料與歷史性(historical)資料 |
| 資料內容 | 作業性資料 | 作業性資料與分析決策用資料 |
| 儲存平台 | 可能為不同平台、分散儲存 | 儲存於單一平台 |
| 資料觀點 | 以作業性或功能性觀點 | 企業主題式資料觀點 |
| 使用觀點 | 使用者不同介面 | 資料不同維度展現 |
| 揮發性 | 資料隨時揮發(更新) | 資料不變性 |

四、針對以下之個體關係圖 (Entity Relationship Diagrams) : (三十分)

- (一)請說明繪製個體關係圖的目的。
- (二)請解釋此圖之意義。
- (三)請說明此圖有何缺點？應如何改善？



答：

(一)個體關係圖為一高階的資料模型，以圖形化的方式表示資料庫概念層的語意。由於概念清楚明瞭，使用者容易接受，因此在資料庫設計上已被廣泛使用。從事資料庫設計時，亦可將資料庫需求先繪製成個體關係圖，再轉為資料庫關聯。

(二)意義：記錄課程、學生、及其上課班級之關係。其中

- 1.課程有其代號、名稱、及學分。
- 2.學生有其代號、時間、及科系。
- 3.上課班級有其代號、時間、及教室。
- 4.每一課程可能有多個上課班級，但每個上課班級只恰屬於一門課程。
- 5.每位學生可以參與多個上課班級，且每個上課班級也可讓多位學生修課。

(三)缺點：

- 1.個體無鍵值屬性，無法唯一辨別記錄。
- 2.無法表達出個體間的關係內容。
- 3.無法明顯表達出上課班級為一課程之弱個體。

可採下面的個體關係圖表示方法，以菱形代表個體間的關係，並以底線註明其鍵值屬性。

