

101年公務人員普通考試試題

代號：41460
44260

全一頁

類 科：統計、資訊處理

科 目：資料處理概要

考試時間：1小時30分

座號：_____

※注意：(一)禁止使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

- 一、假設資料庫運算的兩個交易 τ transaction_1 與 transaction_2，有共用的資料項目 x 跟 y 。也就是說，兩個交易都分別有讀寫共用資料的動作。例如，一個是 {read(x), read(y), write(y), write(x)}，而另一個交易為 {read(y), write(y), read(x), write(x)}。類此的交易運算表示方式，一個包含兩個交易的排程 (schedule)，在什麼條件情況下是不允許並行處理 (concurrent execution)？又在什麼情況下是允許並行處理？除了敘述緣由外，請分別舉例說明。(20分)
- 二、關於資料庫系統的設計，我們常常先定義資料的綱要 (schema)，請問綱要的作用為何？在關聯式資料庫與物件導向式資料庫中，物件導向式資料庫比關聯式資料庫較具優勢之處為何？在考慮關聯式的資料庫設計時，我們常會因正規化的需要，將資料表作分解 (decomposition) 的動作。比如說，一個關於公司人員的資料表，假設該表原包含三個欄位 (名字、職稱、部門)，若資料庫設計者想將其分成兩個表格：(名字、職稱) 與 (職稱、部門)，請問如此分解是好的正規化分解嗎？為什麼？請進一步說明藉由資料表分解，已達到正規化的目的為何？(20分)
- 三、請回答下列問題：(每小題5分，共20分)
- (一)資料結構中，樹 (tree) 的結構與圖 (graph) 的結構有何差別？
- (二)在一個有根節點 (root node) 的樹，其節點 (node) 的數目與邊 (edge) 的數目有何關係？
- (三)在沒方向性的連接圖 (connected graph) 中，一個節點的度數 (degree) 定義為該節點連接的邊的個數。請問其節點的度數與邊的數目有何關係？
- (四)請分別舉出兩個樹結構與圖結構在電腦應用中實際的例子。
- 四、在通訊網路上使用的資料交換技術，以電路交換 (circuit switch) 與封包交換 (packet switch) 為主。請以線路的使用效率與資料傳輸的延遲因素，說明兩種技術間的差異性？在 TCP/IP 網路中，傳輸層的 TCP 與 UDP 的傳輸協定，何者可以稱為電路交換技術的應用？為什麼？(20分)
- 五、假設三種演算法的複雜度函數分別為 $f_1(n) = n^2$, $f_2(n) = 2^n$, $f_3(n) = \log(n!)$ 。將此三個函數依複雜度的高低順序排序，其結果為何？請說明原因？假若 $g(n) = 1 + (1/2) + (1/2^2) + \dots + (1/2^{n-1})$ 。請問“ $g(n) = O(n)$ ”是否正確？為什麼？(註1：此題中 n 為正整數，註2： $O(n)$ 表示 big-O of n 。)(20分)

申論題解答

一、

(一)允許並行處理：

	T1	T2	T3	T4	T5	T6	T7
Transaction_1	Read(x)	Read(y)	Write(y)	Write(x)			
Transaction_2				Read(y)	Write(y)	Read(x)	Write(y)

(二)不允許並行處理：

	T1	T2	T3	T4	T5
Transaction_1	Read(x)		Read(y)		Write(y)
Transaction_2		Read(y)		Write(y)	
				這兩個時間點寫入的資料會不一致	

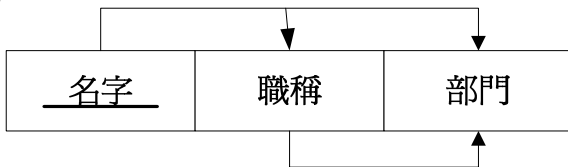
二、

(一)資料綱要就是定義資料儲存在資料庫當中的資料型態和資料名稱。例如我要儲存時間資料，就必須將此欄位的資料型態宣告為 `datetime`；如果要儲存會員帳號，就必須將此欄位的資料型態宣告為 `varchar`

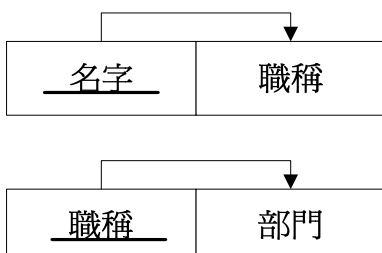
(二)就字面上來看，物件導向資料庫多了物件導向的功能，並同樣使用物件導向的方法模擬傳統資料庫的功能。物件本身可以由其他物件層層包覆所組成，物件與物件間可有各縱向與橫向的繼承關係，同時物件本身亦有其特定的運算式(Method)定義於其上。

應用物件導向的觀念與技術，以定義類別(class)的方式來定義各個資料項目(entity)所具有的欄位。各種資料的運算均以 `function(method)` 定義於類別內部，以 `message passing` 的方式，引發物件去執行 `message` 中所指定的 `method`。

(三)



由上圖可知，此資料表存在遞移相依性，不符合 3NF。因此必須將其拆解為兩個資料表，如下圖：



(四)資料庫正規化的目的在於消除重複的資料欄位，確保資料的一致性。

第一正規化：消除重複資料群，定義出唯一的索引值。

第二正規化：每個非主鍵屬性皆函數相依於主鍵，稱為第二正規化。

第三正規化：當資料表處於第二正規化且不含遞移相依性時，稱為第三正規化。

三、

(一) Tree 是由大於等於 1(不可以為 null)有限個資料所形成之集合。其必須滿足：

1.至少有一個節點，稱之為 root(樹根)。

2.其餘的 Nodes 分成 n 個互斥之集合，T1、T2、T3...Tn 稱為 root 的子樹

Graph 是由 $G=\langle V, E \rangle$ 所組成，其中 V 頂點(Vertex)之集合，E 為邊(Edge)之集合。其種類分為無向圖和有向圖。

(二)節點數(Nodes) = edge + 1

(三)假設一個無向圖具有 n 個節點，則 edge 數目： $n(n-1)/2$

(四)樹狀結構的電腦應用：電腦中的檔案總管，通常它由起始節點開始，連結到一個或數個子項目上，再由子項目繼續連結到一個或數個新的子項目上

圖形結構的電腦應用：最短路徑搜尋

四、

(一)電路交換：我們普遍用的電話網路就是用電路交換 (Circuit Switched) 技術的代表例子。當甲方打電話給乙方時，各地電信局中的交換機會自動找出並建立一條連接甲、乙方的通道，如此資料 (即聲音) 即可透過該條線路直接傳遞，而網路上的節點也能共用有限的線路來傳遞資料。

封包交換：電路交換的缺點是連線建立後會產生「佔線」的問題，而為了解決這個問題，人們又發展出封包交換 (Packet Switched) 的技術：也就是當資料傳遞時會被分割為許多固定長度的封包再送出。如此每一線路皆可傳遞來自不同吏用者的封包，即可解決電路交換方式中的「佔線」問題，並可提高線路使用率。

(二) TCP 與電路交換皆是屬於連接導向的傳輸協定，因此 TCP 可稱為電路交換技術的應用。

五、

(一) $\log(n!) < n^2 < 2^n$

(二) $G(n)=1+(1/2)+(1/2^2)+\dots+(1/2^{n-1})=1+\sum_{k=1}^{n-1} \frac{1}{2^k}=O(1)$

因此 $g(n)=O(n)$ 錯誤