

等 別：高等考試

類 科：資訊技師

科 目：資料結構與資料庫及資料探勘

考試時間：2 小時

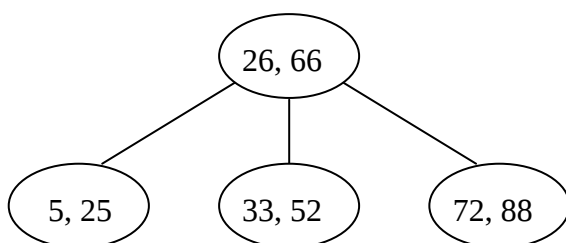
座號：_____

※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、(一)下列圖形是一棵 B-tree of order 3，若陸續加入資料：99、20，會成為怎樣的 B-tree，請畫出最後的 B-tree。(5 分)

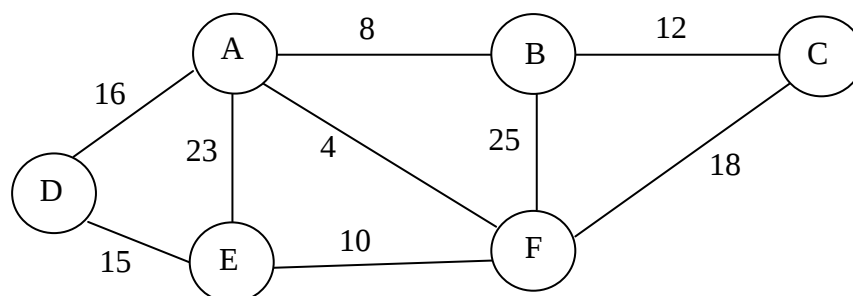
(二)請問 B-tree of order 3 又稱為什麼樹？(5 分)



二、堆疊 (stack) 可應用於後序表示式 (postfix expression) 的運算處理，請利用下列的表示式，繪出堆疊內的變化來說明如何利用堆疊計算其結果，並請寫出演算法。
後序表示式：4 8-9 3 / * (該表示式中的數值均為個位數)。(10 分)

三、請將 18、25、6、33、9、12、55、14 依快速排序 (Quick sort) 由小排到大，若是以最左的資料 18 為支柱點 (pivot)，請繪出 first pass (即是指將資料 18 放在正確的位置) 之排序過程。(5 分)

四、(一)請問什麼是最小成本擴展樹 (minimum cost spanning tree)？(5 分)(二)依據下列的圖形，利用 Prim's algorithm 求出最小成本擴展樹 T，假設 TV 為 T 的頂點集合，設定 TV 的初始值為 A，即 $TV = \{A\}$ 。請繪出最小成本擴展樹的形成過程。(10 分)



五、資料庫管理系統一個重要的工作即是進行交易 (transaction) 管理，在同作控制 (concurrency control) 處理中要確保交易的四項特性，簡稱為 ACID，請逐一說明這四項性質。(15 分)

(請接背面)

等 別：高等考試
類 科：資訊技師
科 目：資料結構與資料庫及資料探勘

六、有一資料表 PRODUCT 之關聯網要 (relation schema) 如下，其中產品代碼為主鍵 (primary key)，除了主鍵之相依性外，若以 $\rightarrow$ 表示相依性 (functional dependency)，該表還存在著右列的相依性，供應商代碼 $\rightarrow$ 供應商名稱

產品代碼	產品名稱	供應商代碼	供應商名稱	價格
------	------	-------	-------	----

請問該資料表符合第三正規化的格式 (third normal form) 嗎？若不符合，請問該如何修改使之符合第三正規化之格式。(10 分)

七、某大學進行全校的體能競賽，下列資料表 GAMERECORD 為參賽選手的分數紀錄範例，該表為關聯資料庫 (relational database) 的資料表 (relation)，一位學生只有一個分數，同一個科系會有多位學生參加，每個學生的學號均不同。請根據該資料表，回答下列第(一)到第(三)題：

GAMERECORD		
STU_NUM (學號)	DEPT_ID (所屬科系代碼)	SCORE (體能分數)
0188561	M1125	80
0366034	S1024	80
0288306	M1125	65

- (一) 請問該資料表的主鍵 (primary key) 應設為該資料表的那些屬性 (attribute)？請說明你的答案。(5 分)
- (二) 寫出“列出平均體能分數大於 80 的科系代碼與該科系的平均體能分數”的 SQL 指令。(5 分)
- (三) 寫出“列出每科系的最高體能分數與科系代碼”的 SQL 指令。(5 分)

八、在資料探勘的分類方法 (cluster analysis) 中有一方法稱為 nearest-neighbor clustering algorithm，請說明該方法的運作原理？(10 分)

九、資料探勘中的關聯法則 (Association rule) 可用來尋找顧客購買產品的關聯規則，{麵包} $\rightarrow$ {牛奶} 表示購買麵包的顧客也會購買牛奶。有一交易紀錄如下表，請寫出關聯規則 {b, c} $\rightarrow$ {d} 的支持度 (support) 和信心度 (confidence)，及其計算過程 (5 分)，並說明這兩者是指什麼？(5 分)

交易代碼	購買的產品
1	{a, b}
2	{a, c, d, e}
3	{b, c, d, e}
4	{a, b, c, d}
5	{a, b, c, f}