資訊

《資料結構》

試題評析

今年高考資料結構試題，難易度分佈平均，有簡單也有難的題目；但是大部份卻都著重於演算方法上，難度雖不算是很難，卻也過於忽略資料結構部份，如樹結構與圖形結構幾乎都沒有命題，有可能會測不出考生的程度。

第一題是雜湊法基本做法的觀念題，考生要拿到分數並不難。第二題是排序法的實作問題，原本應是很簡單，但觀察不仔細，可能會答非所問；而第二與第三兩小題，兩小題答案是相同的，不知是命題有誤，還是故意試探同學的自信心。第三題是基本的互斥集合的問題，拿分應該很容易。第四題是運算式的轉換與求值計算問題，也是容易拿分的題目。第五題則較難了一些，可以用排列組合的觀念來推導，可能會有許多同學臨場反應不過來，而拿不到分數。

綜觀今年高考試題方向，多少有點過於偏頗，但考生若有較完全之準備，想拿到70分以上當不至於太難。而第五題則是關鍵題，能夠正確解答，就有機會取得90分以上的分數，但是第五題卻需要一些排列組合的基礎，才能正確回答，對考生而言是有些困難，不過也正好可以測試考生，平時是否具有以遞迴關係式來求解時間複雜度問題的相關能力。

一、給予如下三個碰撞處理機制(collision handling mechanisms)，請先別描述其原理，再比較其異同。（二十分）

(一)線性探測法(linear probing)

(二)二次式探測法(quadratic probing)

(三)串連法(chaining)

答：

(一) 線性探測法：當碰撞而且溢位發生時，在雜湊表中找尋另一個可用空間的方法，其找尋的方式是由碰撞發生的位置開始，向下循序搜尋下去，遇到第一個有空間的 bucket 時，即使用它來存放資料；若搜尋到表格尾部，可以循環從表格開頭繼續搜尋；若一個循環之後仍未找到空間，則表示整個表格全部滿了，無法再存入任何資料。

(二) 二次探測法：也是當碰撞而且溢位發生時，在雜湊表中找尋另一個可用空間的方法，其找尋的方式是，第一次碰撞時，由碰撞發生的位置加上 1^2；第二次加上，由碰撞發生的位置加上 2^2；第三次加上，由碰撞發生的位置加上 3^2；餘此類推。若超過結尾時，循環由表格開頭繼續下去。

(三) 串連法：當新存入料發生碰撞時，可以在溢位區建立額外的節點來存放，並將節點加入碰撞的 bucket 的串列中，使每個 bucket 各自有一個連結串列。

比較：

(1) (一) 與 (二) 皆為開放式定址法(Open addressing)，意即在原來的雜湊表中找尋可用空間；而 (三) 則是利用到額外的溢位區空間。

(2) (一) 容易發生基本群聚(Primary Clustering)之現象；而 (二) 不會發生基本群聚，但仍可能有 次級群聚(Secondary Clustering)的現象。

(3) (一) 會檢查每一個可用的空間；而 (二) 則可能會漏掉一些可用空間，而誤認為表格已滿。

二、選擇排序法(Selection sort)的基本觀念是從那一堆尚未排序好的資料中，挑選出一個最小的資料，而將它依序置於排序好的位置上。它的演算法如下：

```
Procedure selection
  for i ← 1 to n-1
     min ← i
     for j ← i+1 to n
        if (A[j]<A[min] ) then min ← j
     end for
     temp ← A[i]; A[i] ← A[j]; A[j] ← temp
  end for
```

9 end

問題：

(一)**演算法的第七行是做兩個位置，**A[i]**與**A[j]**內容的交換，明顯的若兩者的內容相同則可不用交換，那麼第七行應改為：**

if (A[i]≠A[j]) then temp ← A[i]; A[i]← A[j]; A[j]←temp

但是一般的選擇排序演算法都沒有這樣的修改，請問為什麼？請寫出理由。（十分）

(二)**演算法的第五行是做資料的比較。對於**n**個資料而言，最多的比較次數是多少？（五分）**

(三)**同**(二)**，最少的比較次數是多少？（五分）**

答：

(一) **一般沒有加上條件檢查，是因為大部份情況下，資料都不相等，因此讓每次交換之前都做一次檢查，反而會使執行時間變長；而節省下來的，可能只是極少數的相同資料的交換時間，所以可能會得不償失。**

(二) $\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}1=\frac{n(n-1)}{2}$。

(三) **與（二）相同，**$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}1=\frac{n(n-1)}{2}$。

三、給予14**個元素，編號**1　14**，如果這些元素間具有如下的等價關係**(equivalence relation)**：**

R={(1,11),(7,11),(2,12),(12,8),(11,12),(3,13)(4,13),(13,14),(14,9),(5,14),(6,10)}

(一)**試找出所有等價類別**(equivalence classes)。**（四分）**

(二)**試提出集合**(set)**表現方式，並定義相關運算，以找出所有等價類別。（十六分）**

答：

(一) {1,2,7,8,11,12}, {3,4,5,9,13,14}, {6,10}。

(二) **因為每一元素只會屬於一個等價類別，所以正好可以用** disjoint sets **的** tree **表示法。**

定義每個元素為一個 tree node**，節點中只要記錄其** parent **為那一個節點；每一棵樹代表一個** disjoint set**，以其** root **來代表。所需要的運算有兩個，即：**

(1) Find(x)**：由元素** x **找尋其所屬的** tree **之** root。

(2) Union(x,y)**：將兩棵樹**(roots **分別為** x **與** y)**合併為一，即讓** x **成為** y **的** parent**，或** y **成為** x **的** parent**，可以用** weighting rule **來判定，何者做為整體的** root **會較佳。**

如此，等價類別的處理方式，是在輸入一個元素對(a,b)**時，先分別找出** a **與** b **所屬的** tree root**，若為同一個，則表示** a **與** b **已合併在同一棵樹；若兩個** roots **不相同，則將兩棵樹合併。最後屬於同一棵樹的元素，則屬於同一個等價類別。**

四、(一)**將一般算術陳式**(expression)**轉換成後序式**(postfix)**有什麼優點？（五分）**

(二)**試將**A+B×C／(E-2)×F**轉換成後序式。（五分）**

(三)**再說明我們如何去計算**(二)**中的後序式，請寫出計算過程。（十分）**

答：

(一) **後序式的點如下：**

(1) **不用知道運算子的優先次序與結合律，即可計算式子之結果，即由左而右先出現先計算。**

(2) **不需要使用括號來改變運算子的計算順序，即可表示任意的計算順序。**

(3) **在電腦內部比中序式容易處理。**

(二) ABC*E2-/F*+

(三) **順序如下：**

(1) A, B, C : push **到** stack

(2) * : **由** stack pop **出來兩個資料，相乘後** push **回** stack

(3) E, 2 : push **到** stack

(4) - : **由** stack pop **出來兩個資料，相減後** push **回** stack

(5) / : **由** stack pop **出來兩個資料，相除後** push **回** stack

(6) F : push **到** stack

(7) * : **由** stack pop **出來兩個資料，相乘後** push **回** stack

(8) +：**由** stack pop **出來兩個資料，相加後** push **回** stack
(9) **最後，由** stack **取出結果即可**

五、某程式的輸入為一串n個介於1到n的正整數 a(1), a(2),..., a(n);**且**a(i)<>a(j) if i<>j。**故共有**n!**種不同輸入，假設每種輸入的發生機率都是一樣。**
給定a(1), a(2),...,a(n),
假設該程式的執行時間經分析後為│a(1)-a(n)│。
令T(n)**為該程式的平均執行時間。**
試列出T(n)**和**T(n+1)**的關係，並說明計算過程。（二十分）**

答：

(一) **當** n=2 **時，**a(1) **與** a(n) **中一定一個是** 1；**另一個是**2，**故** $T(2)=|1-2|=1$，**此為終止狀況。**

而 T(n+1) **可以由排列組合來考慮其與** T(n) **之間的關係：**

(1) **由** 1~n+1 **中，任選其中兩個做為** a(1) **與** a(n+1) **的組合數為** C_2^{n+1}。

(2) **若** a(1) **與** a(n+1) **其中有一個** n+1，**另一個為** 1~n **中某一個的組合數為** C_1^n。

故 a(1) **與** a(n+1) **兩項中，有一項為** n+1，**其機率為** $\dfrac{C_1^n}{C_2^{n+1}}=\dfrac{n}{\frac{n(n+1)}{2}}=\dfrac{2}{n+1}$。**此種狀況之下，另一項可以是** 1~n **中任何一個，故平均值為** $\dfrac{n+(n-1)+...+3+2+1}{n}=\dfrac{\frac{n(n+1)}{2}}{n}=\dfrac{n+1}{2}$。

(3) **若** a(1) **與** a(n+1) **兩項為** 1~n **中某兩個，而皆非** n+1，**則其組合數為** C_2^n，**其機率為** $\dfrac{C_2^n}{C_2^{n+1}}=\dfrac{\frac{n(n-1)}{2}}{\frac{(n+1)n}{2}}=\dfrac{n-1}{n+1}$。**此種狀況之下的平均值，即為** *T(n)*。

綜合上述狀況，可得遞迴關係式

$$T(n+1)=\begin{cases}1 & ,n=1\\ \dfrac{2}{n+1}\times\dfrac{n+1}{2}+\dfrac{n-1}{n+1}\times T(n) & ,n>1\end{cases}$$

化簡之後為

$$T(n+1)=\begin{cases}1 & ,n=1\\ \dfrac{n-1}{n+1}\times T(n)+1 & ,n>1\end{cases}$$

(二) **上面所得之關係式是正確的，但若要藉以求出** T(n) **的公式，較為複雜。下面則是直接以組合數計算** T(n)，**然後可以得到更簡化之結果。**

若 i **與** j **是在** a(1) **與** a(n) **上的兩個數，且假設** i<j，**根據組合可推算**

$$T(n)=\frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-i)}{C_2^n}=\frac{\sum_{i=1}^{n-1}\frac{i(i+1)}{2}}{\frac{n(n-1)}{2}}=\frac{\frac{(n-1)n(n+1)}{6}}{\frac{n(n-1)}{2}}=\frac{n+1}{3}$$，**其中** n>1

而 $T(n+1)=\dfrac{n-1}{n+1}\times T(n)+1=\dfrac{n-1}{n+1}\times\dfrac{n+1}{3}+1=\dfrac{n-1}{3}+1=\dfrac{n+2}{3}$

故 $T(n+1)-T(n)=\dfrac{n+2}{3}-\dfrac{n+1}{3}=\dfrac{1}{3}$

因此可以得到更簡化的遞迴關係式，如下

$$T(n+1)=\begin{cases}1 & ,n=1\\ T(n)+\dfrac{1}{3} & ,n>1\end{cases}$$