

填充題

1. (5%) 中文內碼有許多編碼方式，假設今有一中文碼由兩個位元組所構成。第一個位元組的編碼範圍落在 40~7E 以及 A1~FE 之間，第二個位元組的編碼範圍落在 A0~FE 之間。請問此一編碼最多可表達幾個字？_____
2. (5%) 對於二進位數，其負值我們以二的補數 (2's complement) 計算。請問 10 進位數 23 之 9 的補數是：_____，及它的 10 的補數是：_____。
3. (5%) 執行底下的 C 程式後
X = 50; Y = 0;
for (i = 1; i <= 10; i++)
Y += i;
Z = (X > Y) ? X + Y : X - Y;
請問變數 Z 的值為何？_____

選擇題

4. (5%) 下列有關防火牆之敘述何者有誤？
A、防火牆無法防止內賊對內的侵害，根據經驗，許多入侵或犯罪行為都是自己人或熟知內部網路佈局的人做的；
B、防火牆基本上只管制封包的流向，無法偵測出外界假造的封包，任何人皆可製造假的來源住址的封包；
C、防火牆無法確保連線的可信度，一旦連線涉及外界公眾網路，極有可能被竊聽或劫奪，除非連線另行加密保護；
D、防火牆可以防止病毒的入侵。
5. (5%) 下列有關浮點運算之敘述何者正確？(1) 加、減法的浮點運算會損失數值的精確度；(2) 加、減法的浮點運算不會損失數值的精確度；(3) 乘、除法的浮點運算會損失數值的精確度；(4) 乘、除法的浮點運算不會損失數值的精確度。
A、敘述 1、4 正確
B、敘述 2、3 正確
C、敘述 1、3 正確
D、敘述 2、4 正確
6. (5%) 資料傳輸時可作雙向傳輸，但無法同時雙向傳輸的傳輸方法為
A、單工
B、半雙工
C、多工
D、全雙工
7. (5%) 下列之敘述何者正確？(1) 有些計算機會將一部份軟體存於唯讀記憶體中，稱之為輔助記憶體；(2) CPU 之讀、寫、中斷及重置等信號乃是由控制匯流排來傳送；(3) 暫存器 (Register) 之數目多寡為影響顯示幕解析度的因素之一；(4) 一個硬碟可同時安裝多種作業系統，但同一時間只允許使用一個作業系統。
A、敘述 1、4 正確
B、敘述 2、3 正確
C、敘述 1、3 正確
D、敘述 2、4 正確

8. (5%) 下列關於程式語言之敘述何者有誤？
- A、C 是 procedural language
 - B、C 是 strong-typing language
 - C、Java script 是 Java 的 sub-language
 - D、以上皆是
9. (5%) 下列有關快取記憶體 (cache memory) 之敘述何者正確？(1) 當區塊大小 (block size) 由很小開始加大時，快取命中率 (hit rate) 會增加；(2) 當區塊大小由很小開始加大時，快取命中率會降低；(3) 當區塊大小已經夠大仍繼續增加時，快取命中率會增加；(4) 當區塊大小已經夠大仍繼續增加時，快取命中率會降低。
- A、敘述 1、4 正確
 - B、敘述 2、3 正確
 - C、敘述 1、3 正確
 - D、敘述 2、4 正確
10. (5%) 下列關於 POS 系統之敘述何者正確？
- A、POS 系統用在電子商務的採購管理
 - B、POS 系統用在零售業的銷售管理
 - C、POS 系統是 ERP 系統的一個模組
 - D、以上皆是

簡答題

11. (10%) Write a pseudo code to exchange the row and column elements of a matrix.
12. (10%) Write a pseudo code to find the longest common substring of two strings.
13. Answer each of the following problems
- (a) (4%) Distinguish between fixed partitions and dynamic partitions in a memory management system.
 - (b) (9%) In a dynamic partition memory management system, if the current value of the base register is 42993 and the current value of the bounds register is 2031, compute the physical addresses that correspond to the following logical addresses: (1) 104, (2) 1755, (3) 3041
14. Answer each of the following problems
- (a) (5%) Use quick sort to determine the order of the records with keys (37, 1, 5, 26, 12, 60, 72, 10, 10, 23). Describe the sorting steps and result.
 - (b) (3%) Is quick sort a stable sorting method? Explain why?
 - (c) (5%) Use quick sort to determine the order of the records with keys (5, 4, 1, 8, 3, 2, 7). Describe the sorting steps and result.
 - (d) (4%) What kind of data structure you would use for the heap in the heap sort program? Explain why.

1. 在資訊科技的演進與推動下，形成了一個迥異於傳統經營模式的經濟體系。在此體系下，有所謂的(1)數位化，(2)資訊化(電腦化)，(3)電子化(e化)，(4)虛擬化，(5)知識化，(6)網路化，請說明上述六個演變各自不同的強調重點及其對企業經營模式的影響?(35分)
2. 在研究使用者接不接受新資訊系統的引進時，最著名的一個理論稱之為科技接受模式(Technology Acceptance Model, TAM)，請繪圖並說明 TAM 的意涵。(15分)
3. 在實體的商務模式中，由於店面、陳列空間與動線的限制，商品的陳列常以熱門商品為優先，因此 80/20 法則被許多以傳統商業模式為主的企業或商家奉為圭臬。該法則之概念為一個企業的利潤有 80%是來自於 20%的熱門商品。然而，長尾理論 (Theory of the Long Tail) 於 2004 年 10 月在 *Wired* 雜誌發表後，對電子商務的商業模式帶來不同的經濟思維。請以網際網路電子商務之 Amazon.com 商店為例，回答下列問題：(25分)
 - (1) 請說明長尾理論的概念。(5分)
 - (2) 在 Amazon.com，長尾指的是什麼?(5分)
 - (3) 在 Amazon.com，資訊科技如何實現長尾理論?(15分)
4. 由於網路與資訊科技的進步、普及與創新應用，組織不但可藉此改善生產率，也可能對其客戶提供新的服務。請針對客製化產品的訂購流程，說明組織如何應用網路與資訊科技對客戶提供 Mass Customization 之服務。(25分)

請依照題號書寫答案，並標示清楚題號。下列表格為機率分佈之變數與機率值。

α	.005	.01	0.025	0.05	0.1
z_α	2.58	2.33	1.96	1.645	1.28
$\chi_\alpha^2(1)$	7.88	6.63	5.02	3.84	2.71
$\chi_\alpha^2(2)$	10.60	9.21	7.38	5.99	4.61
$\chi_\alpha^2(3)$	12.84	11.34	9.35	7.81	6.25
$\chi_\alpha^2(4)$	14.86	13.28	11.14	9.49	7.78

Note: $\alpha = P(z > z_\alpha)$

選擇題 (1 至 10 題，每題 5 分)：

- A consumer group wants to estimate the proportion of university students that have part-time jobs. Within the error of 4% with 95% confidence, how large the sample size is required?

(a) 1068 (b) 1113 (c) 846 (d) 423 (e) 601
- Jason claims that he can correctly identify whether a man is married or single by the way he dresses. Suppose in actuality that if a man is married, Jason can correctly identify the man as married 75% of the time. When a man is single, Jason will incorrectly identify that man as married 20% of the time. Presented with a man and asked to identify the marital status of this man (who is either married or single), Jason considers this to be a hypothesis test with the null hypothesis being that the man is married and the alternative that the man is single. What would be a Type I error?

(a) Saying that the man is married when in fact the man is married.
 (b) Saying that the man is married when in fact the man is single.
 (c) Saying that the man is single when in fact the man is married.
 (d) Saying that the man is single when in fact the man is single.
 (e) None of the above.
- Which of the following is wrong for the normal distribution?

(a) During the range of plus and minus one standard deviation around the mean contains almost the probability of 0.5.
 (b) During the range of plus and minus two standard deviations around the mean contains almost the probability of 0.95.
 (c) Its measures of central tendency (mean, median, and mode) are all identical.
 (d) Its associated random variable has an infinite range ($-\infty < X < \infty$)
 (e) A normal distribution does not show as M-shape with two peaks.

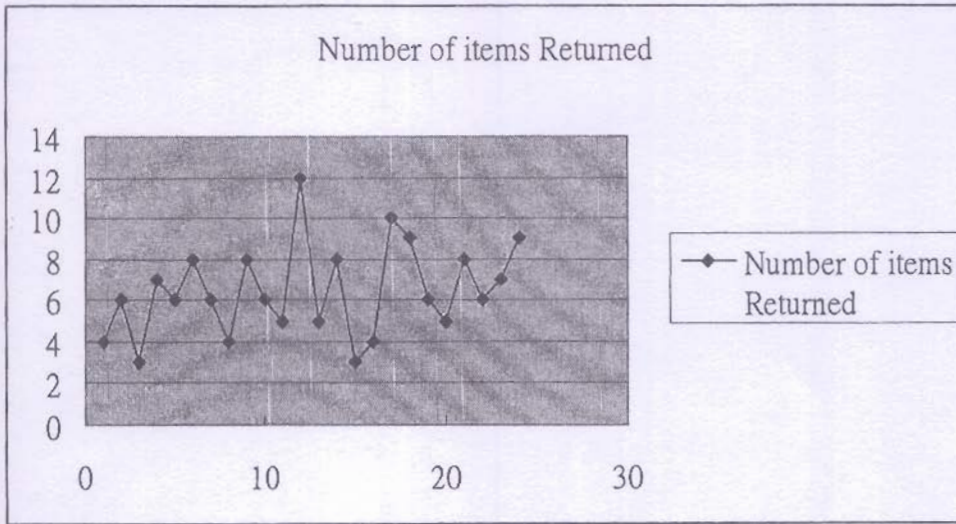
4. An university randomly selects 200 students and finds that 65% have learned a second foreign language. The university did a similar study 10 years ago in which 30% of a random sample of 180 students had learned a second foreign language. Let π_1 and π_2 represent the true proportion of workers who have learned a second foreign language in the recent study and the past study, respectively. If the university wanted to test whether this proportion has be increased from the previous study, with the significance level $\alpha = .05$, which of the following is most correct?
- (a) $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$; reject the null hypothesis and conclude that the proportion of students who have learned a second foreign language has changed over the intervening 10 years.
 - (b) $H_0 : \pi_1 - \pi_2 \leq 0$ versus $H_1 : \pi_1 - \pi_2 > 0$; reject the null hypothesis and conclude that the proportion of students who have learned a second foreign language has increased over the intervening 10 years.
 - (c) $H_0 : \pi_1 - \pi_2 \geq 0$ versus $H_1 : \pi_1 - \pi_2 < 0$; do not reject the null hypothesis and conclude that the proportion of students who have learned a second foreign language has increased over the intervening 10 years.
 - (d) $H_0 : \pi_1 - \pi_2 \leq 0$ versus $H_1 : \pi_1 - \pi_2 > 0$; do not reject the null hypothesis and conclude that the proportion of students who have learned a second foreign language has actually decreased over the intervening 10 years.
 - (e) $H_0 : \pi_1 - \pi_2 = 0$ versus $H_1 : \pi_1 - \pi_2 \neq 0$; do not reject the null hypothesis and there is insufficient evidence to conclude the proportion of students who have learned a second foreign language has changed over the intervening 10 years.
5. Recent studies have found that children are more obsess than in the past. The amount of time children spent watching television has received much of the blame. A survey of 500 eight-year-olds revealed the following with regards to weights and average number of hours a day spent watching television. We are interested in testing whether the average number of hours a day spent watching television and weights are independent at 1 % level of significance.

Average time spend reading story books	Average time spend watching TV	
	Less than 2 hours	More than 2 hours
Less than 2 hours	90	230
More than 2 hours	22	158

In order to approach the problem, which of the following is most correct?

- (a) Use Z-test. The null hypothesis is the average TV watching time is more or equal to the average story-book reading time. Reject the hypothesis.
- (b) Use Z-test. The null hypothesis is the average TV watching time is more or equal to the average story-book reading time. Do not reject the hypothesis.
- (c) Use χ^2 -test. The null hypothesis is there is no connection between TV watching time and story-book reading time. Reject the hypothesis.
- (d) Use χ^2 -test. The null hypothesis is there is no connection between TV watching time and story-book reading time. Do not reject the hypothesis.
- (e) Use ANOVA table to analyze the data. The conclusion is to reject the hypothesis.

6. To improve the service quality, the owner of a dry cleaning business wants to study the number of dry-cleaned items that are returned for rework per day. Records were shown in the figure below.



For the data, the control limits are given as (Lower Control Limit, Average returned items per day, Upper Control Limit) = (0, 6.46, 14.08). Which of the following is incorrect? (Note that $UCL = 14.08 = 6.46 + 3 * \sqrt{6.46}$)

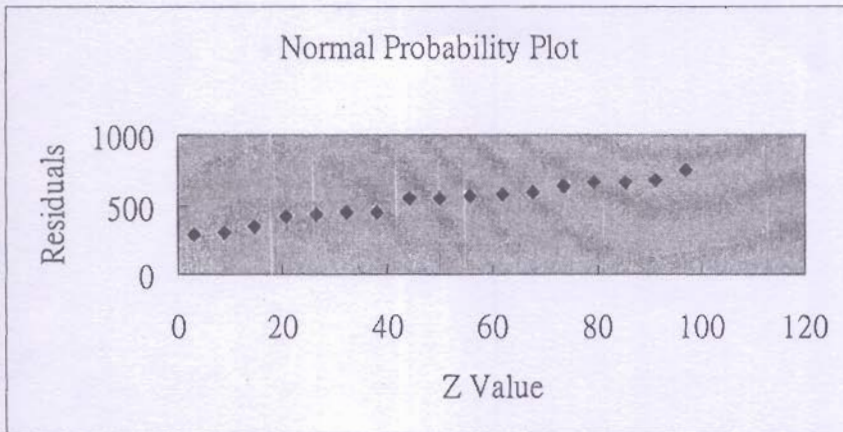
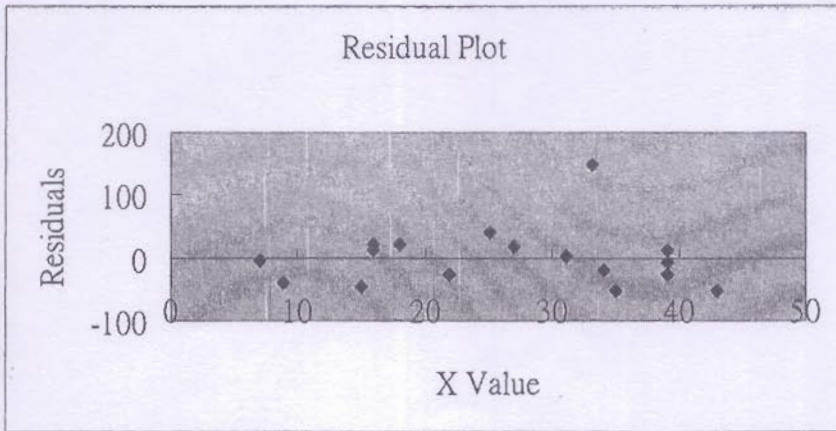
- (a) The process is in a state of statistical control.
- (b) The returned items more than or less than the average returned item value 6.46 shows not pattern. That is, the signaling shows no trend.
- (c) Although the sample size is not big, the owner is able to refer the data to reduce the number of items that are returned for rework.
- (d) The peak happens on day 12 with 12 items returned. The incidence happens occasionally. The manager should ignore it.
- (e) This is an control chart approach.

題組：第 7 題至第 8 題

Following is an output from a simple linear regression along with the residual plot and normal probability plot obtained from a data set of 17 pairs of $(x_i, y_i), i = 1, 2, \dots, 30$.

ANOVA

	DF	SS	MS	F	Significance F
Regression	1	265769.1	265769.1	108.1697	2.97E-08
Residual	15	36854.48	2456.965		
Total	16	302623.5			



7. Referring to the output, what is the standard error of estimate?

- (a) 2456.965
- (b) 36854.48
- (c) 265769.1
- (d) $302623.5/16 = 18913.97$
- (e) 1.090553

8. Referring to the output, which of the following statement is false?

- (a) The normality or error assumption appears to be hold.
- (b) The errors appear to be random.
- (c) The homoscedasticity of error assumption appears to be hold.
- (d) There appears to be autocorrelation in the residuals
- (e) The regression effect is statistically significant.

題組：第 9 題至第 10 題

An advertising agency plans to test the effect of internet advertising on product perception. An experiment is designed to compare six different internet advertisements, A, B, C, D, E, F, each with different advertising approaches. The agency asks for 7 students to watch through all six advertisements and then rate each of them with ranges from 1 to 20. The scores for the 7 students are as follows.

		Internet Advertisement					
		A	B	C	D	E	F
Students	1	17	17	7	6	12	9
	2	18	16	8	9	19	7
	3	15	16	10	13	18	11
	4	19	20	15	11	12	13
	5	20	19	14	9	17	10
	6	19	17	13	5	14	15
	7	16	15	7	7	12	8

Complete the following ANOVA Table and fill in the number

ANOVA					
Source	SS	Degrees of Freedom	MS	F	P-value
Among Treatments	522.79	DF1	MS1	F1	1.65E-08
Among Blocks	95.57	DF2	MS2	F2	0.0237
Error	164.71	DF3	MS3		
Total	783.07				

9. Degrees of freedom are (as "DF1, DF2, DF3" shown in the ANOVA Table)

- (a) 6, 5, 30 (b) 5, 36, 41 (c) 6, 5, 36 (d) 5, 6, 36 (e) 5, 6, 30

10. Give the significance level of 0.01, which of the following is correct?

- a) There is evidence of a difference among the students.
- b) There is evidence of a difference among the internet advertisements.
- c) The blocking has a significant effect in reducing the random error.
- d) There is evidence of differences among students and among advertisements.
- e) There is no evidence of differences among students and among advertisements

計算題 (第 11 至 14 題)

11. 假設 10 個商品中有 2 個有缺點，方法一由這 10 個商品中隨機選取 1 個，令 X 為有缺點的商品個數， $P(X)$ 為其機率函數。方法二找出缺點的測試成本為 500 元；方法二由這 10 個商品中隨機選取 4 個，令 Y 為有缺點的商品個數， $P(Y)$ 為其機率函數，方法二找出缺點的測試成本為 1600 元。

- 列出 X 的機率分佈，即 x 與 $f(x)$ 。(5 分)
- 列出 Y 的機率分佈，即 y 與 $f(y)$ 。(5 分)
- 分別計算 X 和 Y 的期望值。(5 分)
- 分別計算方法一和方法二找出一個缺點商品的平均測試成本。(5 分)

12. 令 X 為一 gamma 分佈，其機率密度函數為

$$f(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-x/\beta}, 0 < x < \infty$$

其中 $\Gamma(\alpha) = (\alpha-1)!$ 若 α 為大於 1 的正整數。

- 令 $\alpha=2$, $\beta=4$ 。計算 $x=0, 5, 10, 15, 20$ 時 $f(x)$ 的值。並以 x 為橫軸， $f(x)$ 為縱軸畫出 X 的近似機率分佈。(6 分)
- 計算 $\alpha=1$, $\beta=4$ 條件下 X 的期望值。(5 分)

13. 假設一天 24 小時中某一信箱收到的廣告信件個數和非廣告信件個數均為 Poisson 分佈，平均值分別為 4 封與 2 封。

- 試計算 8 小時的工作時間中總共收到 3 封信件的機率。(5 分)
- 試計算在下班前一個小時內會收到 1 封非廣告信件的機率。(5 分)

註: Poisson 分佈為 $P(X=n) = \frac{e^{-\lambda} \lambda^n}{n!}$

14. 當下列條件變化而其他條件不變的情況下，母體平均數信賴區間的估計會如何變化？

- 樣本數 n 增加。(3 分)
- 信賴水準 $1-\alpha$ 提高。(3 分)
- 母體變異數 σ^2 變大。(3 分)

每小題均回答變大、變小、不變或沒有必然關係。

填充題

1. (5%) 中文內碼有許多編碼方式，假設今有一中文碼由兩個位元組所構成。第一個位元組的編碼範圍落在 40~7E 以及 A1~FE 之間，第二個位元組的編碼範圍落在 A0~FE 之間。請問此一編碼最多可表達幾個字？_____
2. (5%) 對於二進位數，其負值我們以二的補數 (2's complement) 計算。請問 10 進位數 23 之 9 的補數是：_____，及它的 10 的補數是：_____。
3. (5%) 執行底下的 C 程式後
X = 50; Y = 0;
for (i = 1; i <= 10; i++)
Y += i;
Z = (X > Y) ? X + Y : X - Y;
請問變數 Z 的值為何？_____

選擇題

4. (5%) 下列有關防火牆之敘述何者有誤？
A、防火牆無法防止內賊對內的侵害，根據經驗，許多入侵或犯罪行為都是自己人或熟知內部網路佈局的人做的；
B、防火牆基本上只管制封包的流向，無法偵測出外界假造的封包，任何人皆可製造假的來源住址的封包；
C、防火牆無法確保連線的可信度，一旦連線涉及外界公眾網路，極有可能被竊聽或劫奪，除非連線另行加密保護；
D、防火牆可以防止病毒的入侵。
5. (5%) 下列有關浮點運算之敘述何者正確？(1) 加、減法的浮點運算會損失數值的精確度；(2) 加、減法的浮點運算不會損失數值的精確度；(3) 乘、除法的浮點運算會損失數值的精確度；(4) 乘、除法的浮點運算不會損失數值的精確度。
A、敘述 1、4 正確
B、敘述 2、3 正確
C、敘述 1、3 正確
D、敘述 2、4 正確
6. (5%) 資料傳輸時可作雙向傳輸，但無法同時雙向傳輸的傳輸方法為
A、單工
B、半雙工
C、多工
D、全雙工
7. (5%) 下列之敘述何者正確？(1) 有些計算機會將一部份軟體存於唯讀記憶體中，稱之為輔助記憶體；(2) CPU 之讀、寫、中斷及重置等信號乃是由控制匯流排來傳送；(3) 暫存器 (Register) 之數目多寡為影響顯示幕解析度的因素之一；(4) 一個硬碟可同時安裝多種作業系統，但同一時間只允許使用一個作業系統。
A、敘述 1、4 正確
B、敘述 2、3 正確
C、敘述 1、3 正確
D、敘述 2、4 正確
8. (5%) 下列關於程式語言之敘述何者有誤？
A、C 是 procedural language

- B、C 是 strong-typing language
 C、Java script 是 Java 的 sub-language
 D、以上皆是
9. (5%) 下列有關快取記憶體 (cache memory) 之敘述何者正確？(1) 當區塊大小 (block size) 由很小開始加大時，快取命中率 (hit rate) 會增加；(2) 當區塊大小由很小開始加大時，快取命中率會降低；(3) 當區塊大小已經夠大仍繼續增加時，快取命中率會增加；(4) 當區塊大小已經夠大仍繼續增加時，快取命中率會降低。
- A、敘述 1、4 正確
 B、敘述 2、3 正確
 C、敘述 1、3 正確
 D、敘述 2、4 正確
10. (5%) 下列關於 POS 系統之敘述何者正確？
- A、POS 系統用在電子商務的採購管理
 B、POS 系統用在零售業的銷售管理
 C、POS 系統是 ERP 系統的一個模組
 D、以上皆是

簡答題

11. (10%) Write a pseudo code to exchange the row and column elements of a matrix.
12. (10%) Write a pseudo code to find the longest common substring of two strings.
13. Answer each of the following problems
- (a) (4%) Distinguish between fixed partitions and dynamic partitions in a memory management system.
- (b) (9%) In a dynamic partition memory management system, if the current value of the base register is 42993 and the current value of the bounds register is 2031, compute the physical addresses that correspond to the following logical addresses: (1) 104, (2) 1755, (3) 3041
14. Five processes *A*, *B*, *C*, *D* and *E* arrived in this order at the same time with the following CPU burst and priority values (a smaller value means a higher priority).

<i>Process</i>	<i>CPU Burst</i>	<i>Priority</i>
<i>A</i>	4	2
<i>B</i>	6	4
<i>C</i>	2	1
<i>D</i>	5	3
<i>E</i>	3	5

Indicate the waiting time and average waiting time for processes *A*, *B*, *C*, *D* and *E* when each of the following scheduling policies is used (1) First-Come-First-Served; (2) Non-Preemptive Shortest-Job First; (3) Priority; (4) Round-Robin (time quantum=2).

1. (20%)

Consider the following recurrence function:

$$T(n) = \begin{cases} \theta(1) & \text{if } n = 1 \\ 4T(\frac{n}{2}) + \theta(n^2) & n > 1 \end{cases}$$

Suppose I guess $T(n) = O(n^2)$ and prove its correctness by the following deduction:

$$T(n) = 4 \times O\left(\frac{n^2}{2^2}\right) + n^2 = O(n^2) + n^2 = O(n^2)$$

- (A) What's wrong with the above deduction? (5%)
- (B) Please derive the solution of $T(n)$. (不可以套公式) (15%)

2. (20%)

Suppose you have three coins, A, B, and C. A is a fair coin, B has 3/4 probability of coming up with head, and C always comes up with head.

- (A) If you choose B and toss it three times, what is the probability that you get three heads? (5%)
- (B) If you randomly choose a coin, toss it three times, and get three heads, what is the probability that the coin you chose is C? (15%)

3. (10%)

$$f(x) = f\left(\left\lceil \frac{x}{2} \right\rceil\right) + f\left(\left\lfloor \frac{x}{2} \right\rfloor\right) + 1$$

The following C-like program is meant to implement the above recurrence function.

```
int f(int x)
{
    if (x <= 0)
        return 0;
    return (f((x+1)/2)+f(x/2)+1);
}
```

Will this program return a correct value with any integer for x? Please justify.

4. (10%)

Consider a binary tree T , where each internal node stores a single character. If the inorder traversal of T yields GOODBOOK, and the postorder traversal of T

yields GOOGBKOD. What is the preorder traversal of T ?

5. (10%)

Consider the design and implementation of a calendar (行事曆) for the year of 2009. The program will provide a function for the user to add ANY number of memo entries to ANY date whenever the user wishes to add. Note that the amount of days in 2009 is fixed (i.e. 365), while the amount of entries can be added to any day is unlimited. (在 365 任何一個日子裡可加入隨意多的行程)

(A) Which data structure will you choose for the 365 days in the calendar? Why? You are welcomed to draw picture to illustrate your idea and the usage of the data structure. (5%)

(B) Which data structure will you choose for each memo entry in each day of the calendar? Why? You are welcomed to draw picture to illustrate your idea and the usage of the data structure. (5%)

6. (10%)

Consider a hash table which is used to map the scores of 50 students in a classroom to 13 buckets. The key of mapping is the student ID which has 9 numerical digits (學號有九位數).

(A) Please advise a mapping function (hash function) for this hash table. (5%)

(B) Please describe how to resolve the conflict in which more than one student ID is mapped to the same bucket. (5%)

7. (20%)

$x \% y$ in C means the remainder of x divided by y . Make use of it when you solve the following sub-problems.

(A) Formulate GCD (the greatest common divisor) of two integers x and y in recursive form. (10%)

(B) Write a recursive C function to compute $\text{GCD}(x,y)$. (10%)