【不可使用計算機】＊作答前請先核對試題、答案卷（試卷）與准考證之所組別與考科是否相符!!

1. (16%) Explain the following terms

(a) Automatic computing (4%)

(b) Petabyte (4%)

(c) SoLoMo (4%)

(d) Social Search (4%)

2. (9%) Explain the purposes and procedures of database (a) normalization and (b) de-normalization. Please illustrate by examples. (5%+4%)

3. (9%) (a) Explain the characteristics of and (b) describe the corresponding computing techniques/algorithms for "Big Data". (5%+4%)

4. (10 %) For every tree $T = (V, E)$, prove that (a) $|V| = |E| + 1$, and that (b) if $|V| \geq 2$, then T has at least two vertices with degree 1. (5%+5%)

5. (10%) Let $X[1:m]$ and $Y[1:n]$ be strings defined by four symbols A, C, G, and T. For example, $X[1:7]$ = GACTATG and $Y[1:8]$=AGTCGGCA. Given strings $X[1:m]$ and $Y[1:n]$, define a recursive function $f(m, n)$ as follows:

$f(0, 0)=0$;

$f(i, 0)=0,\ 1 \leq i \leq m$;

$f(0, j)=0,\ 1 \leq j \leq n$;

for $1 \leq i \leq m, 1 \leq j \leq n$,

$$f(i,j) = \begin{cases} f(i-1, j-1) + 1, & \text{if } X[i] = Y[j]; \\ \max\{f(i, j-1), f(i-1, j)\}, & \text{otherwise.} \end{cases}$$

(a) Write non-recursive pseudo codes for finding $f(m, n)$. (5%)

(b) Analyze the required running time, in the Big-O notation, of your codes. (5%)

6. (13%) (Stable Marriage Problem) Consider n men and n women for match-making of n pairs of opposite sexes. Each person has a preference list of all members of the opposite sex. The following instance consists of 5 men and 5 women. By his preference list, man 1 ranks woman 2 the highest, woman 3 the second, and so on; in woman 1's preference list, she prefers man 5 the highest and man 2 the least. The problem is to match n disjoint pairs of men and women. A matching is called *stable* whenever it is not a case where a man m of some pair prefers a woman w of another pair and the woman w also prefers the man m more than her current partner.

Men	Preference lists
1	2, 3, 4, 5, 1
2	1, 5, 2, 4, 3
3	5, 1, 3, 4, 2
4	3, 2, 5, 1, 4
5	4, 5, 2, 1, 3

Women	Preference lists
1	5, 4, 1, 3, 2
2	2, 5, 3, 1, 4
3	1, 3, 2, 4, 5
4	4, 2, 5, 3, 1
5	1, 5, 4, 2, 3

(a) Is the matching {(man, woman): (1, 1), (2, 2), (3, 3), (4, 4), (5, 5)} stable? Please give your reasoning. (3%)

(b) Use ALGORITHM STABLE-MARRIAGE given in the following box to derive a stable matching of the above 5-man, 5-woman instance. The algorithm is based on the strategy that men propose to women and women make decisions on acceptance or rejection. (5%)

(c) What is the running time, in the Big-O notation, of the algorithm? (3%)

(d) Suggest applications of the stable marriage problem.(2%)

【不可使用計算機】＊作答前請先核對試題、答案卷（試卷）與准考證之所組別與考科是否相符!!

```
ALGORITHM STABLE-MARRIAGE
set all men and all women free;
while there exists a free man m who has a woman w to propose to
{
    let w be the highest ranked woman to whom man m has not yet proposed
    if w is free, then collect the pair (m, w);
    else /* some pair (m', w) already exists
        if w prefers m more than m'
            collect the pair (m, w) and set m' free
}
```

7. (6%) Evaluate and give *tight* bounds (in the Big-O notation) for the following recurrence functions $T(n)$ where $T(1)=1$:
(a) $T(n) = 2T(n/2)+n$
(b) $T(n) = 2T(n/2)+n^2$

8. (8%) You are given a computer network (directed graph $G = (V,E)$), in which each link/edge e has a probability of *pe* in [0,1] to be alive. That means that at any given time, with probability *pe*, the link is up, and with probability 1 - *pe*, it is down. We assume that each edge is up or down independently of all other edges. You want to send an important packet from s to t along a single path, and therefore want to choose the path P with the highest probability of having all edges in P up simultaneously. Give an algorithm with running time polynomial in $|V|$ and $|E|$ for finding such a path.
Hints: notice the similarity between this problem and the (single-source) shortest path problem, which can be solved by Dijkstra's shortest path algorithm.

9. (8%) Write a SQL query to find the number of employees in each department when you have the following tables in a relational database (where the underlines mark the keys):
Employees(<u>*Emp_ID*</u>,*Emp_Name*,*Dept_ID*)
Departments(<u>*Dept_ID*</u>,*Dept_Name*)

10. (11%). Suppose that you are given complete binary tree T with a root r, and on each node v in T a prize $pv \geq 0$. You are also given a number $k \geq 0$, the number of prizes you are allowed to pick up.
Your goal is to select a subtree of k nodes rooted at r, and containing the largest total prize. Notice that the "subtree" means that your selected set of nodes must be connected (you can't skip some nodes).
Here is a reference relation for this problem if you are using dynamic programming: in the description, we assume that v is a node, and unless it is a leaf, its children are $v1$ and $v2$.
$\text{OPT}(0,v) = 0$, for all nodes v
$\text{OPT}(i+1,v) = pv$, for all leaves v and all i
$\text{OPT}(i+1,v) = pv + \max_{0\leq j\leq i}(\text{OPT}(j,v1) + \text{OPT}(i-j,v2))$, for all internal nodes v and all i.

(a) (3%)Explain what the meaning of $\text{OPT}(i,v)$ is here.
(b) (6%) Prove that recurrence above correctly captures the meaning. (base cases? explanation of "+1", etc.)
(c) (2%) How would you get the best overall prize (once you have computed $\text{OPT}(i,v)$ for all i and v)?