THE UNIVERSITY OF HONG KONG

FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE

COMP3270 ARTIFICIAL INTELLIGENCE

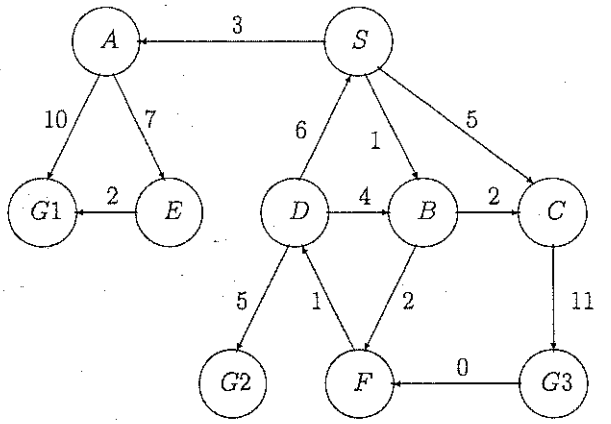
Date: 21 Dec 2017 Time: 9:30am – 12:30pm

Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is the candidate's responsibility to ensure that their calculator operates satisfactorily, and candidates must record the name and type of the calculator used on the front page of the examination script.

Instructions:

- This paper contains 6 questions on 6 pages.
- The mark for each question is enclosed in a pair of square brackets after the questions.
- Answer ALL questions.
- Total Mark is 100.

1. Consider the following search space, where S is the start node, and $G1$, $G2$ and $G3$ are goal states. Search with multiple goals proceeds the same way as search with a single goal, except that the algorithm terminates when any of the goal states is reached by an optimal path. The links are labeled with the cost of traversing them and the heuristic cost to a goal is reported in the table below.



node	S	A	B	C	D	E	F	G1	G2	G3
h	8	9	1	3	4	1	5	0	0	0

(a) Apply *Best-first Search* algorithm, and report the resulting path and its path cost. [2]

(b) Apply the A^* algorithm on the above graph, filling in the table with (f, g, h) after the node name. The first two lines are filled as example. If you find a new path to a node that is already on the queue, update its cost (using the lower f value) instead of adding another copy of that node to the queue. Keep the queue sorted. [6]

Iteration	Node Expanded	Sorted Queue, $Node(f, g, h)$
0	–	S(8,0,8)
1	S	A(12,3,9), B(2,1,1), C(8,5,3)
2	?	...

(c) For the solution found by A^*, give the cost and sequence of nodes comprising the path. [2]

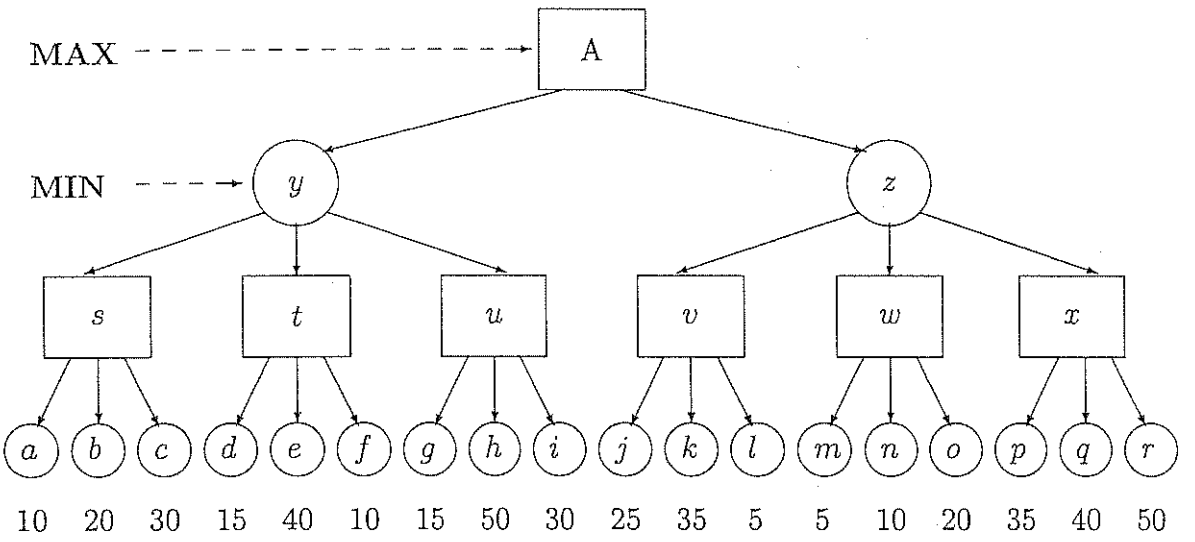
(d) For each of the following algorithms, state whether it is guaranteed to converge to the *global* maximum. Assume that the state space is finite. Justify your answer in one sentence for each algorithm. [6]

(i) Hill-climbing from a randomly chosen initial condition.

(ii) Simulated annealing

(iii) Genetic algorithm

2. Given the following game tree, where a rectangle represents a max node, and a circle represents a min node. The evaluation of the nodes of the last row is provided.



(a) Apply minimax procedure to this game tree. Write down the evaluation of all nodes in the tree. What should be the next move of A. [5]

(b) Apply α-β pruning to this game tree. Draw the resultant search tree that is built by this algorithm. Remember to show the place where α-cut or β-cut occurs. [7]

(c) The above game tree applies to deterministic games. However, there are games that involve chances.

(i) Give an example of chances in a game. [1]

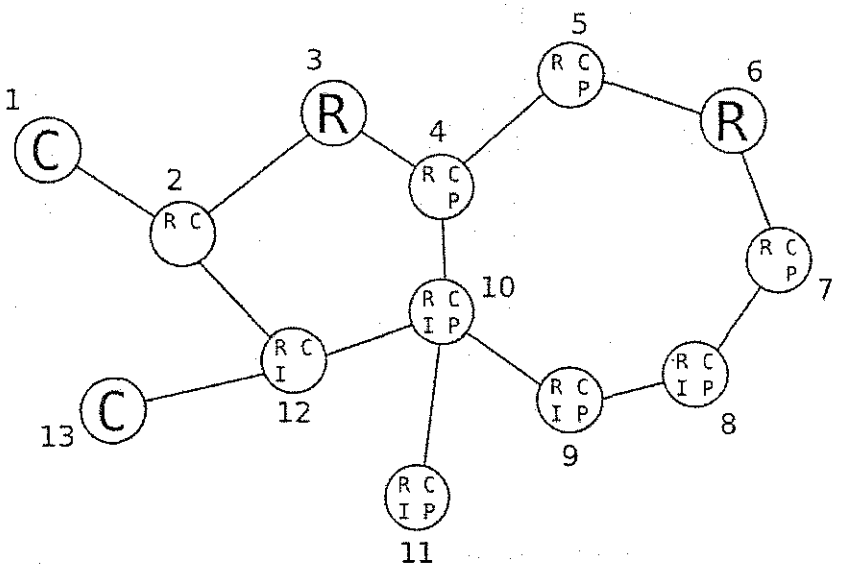
(ii) How can you modify the game tree to incorporate chances. [2]

(iii) It is known that as long as the order of evaluation function is preserved, minimax procedure will produce the same result. Is it also true for games with chance? Briefly explain. [3]

3. (a) City Hong is planning a development project for a new town. The new town is divided into 13 regions and each region can be of the type *Residential*(**R**), *Commercial*(**C**), *Industrial*(**I**) or *Park*(**P**), and the following rule applies:

1. A **Residential** region can only be adjacent to another **Residential** region, a Commerical Region or a **Park**.
2. A **Commerical** region can only be connected to a **Residential** region, another **Commercial** region, or an **Idustrial** region.
3. An **Industrial** region can only be adjacent to a **Commercial** region, or another **Industrial** region.
4. A **Park** can only be connected to a **Residential** region.

The Urban Planning department decide to use constraint propagation to design the new town, and starts with the following partially completed plan.



(i) In solving CSP, we use *forward checking* and *arc consistency* to reduce the domain size. Explain what they are. [5]

(ii) If the urban planner decided to assign region 10 to be an **Industrial** region. Using forward checking only, cross out values in the domain of the variables affected in the above diagram. [3]

(iii) What will be the resulting domain if both *forward checking* and **arc consistency** are used. [5]

(b) Formulate the following cryptarithmetic puzzle as a constraint satisfaction problem. State what are the variables, the domain and the constraints. [4] (You are not required to solve the puzzle)

```
   SEND
+) MORE
-------
  MONEY
```

4. Peter, Paul and Mary belong to the Hoofers Club. Consider the following sentences:

   1. Every member of the Hoofers Club is either a skier or a mountain climber or both.
   2. No mountain climber likes rain.
   3. All skiers like snow.
   4. Mary dislikes whatever Peter likes and likes whatever Peter dislikes.
   5. Peter likes rain and snow

   Using the following symbol *only*:

`S(x)`	`x` is a skier
`M(x)`	`x` is a mountain climber
`L(x,y)`	`x` likes `y`

   where the domain of `x` is members of Hoofers Club, and domain of `y` is rain and snow.

   You are going to prove that "Mary is a mountain climber but not a skier".

   (a) Translate the above sentences and the query to first order logic sentences. [6]

   (b) Convert the sentences to Conjunctive Normal Form. [6]

   (c) Apply resolution with suitable substitutions to show that the query "Mary is a mountain climber but not a skier" is true. [6]

5. (a) Given the following joint distribution:

	toothache		¬ toothache	
	catch	¬ catch	catch	¬ catch
cavity	0.15	0.03	0.12	0.04
¬ cavity	0.03	0.07	0.14	0.42

Find the probability distribution $\mathbf{P}(Cavity|toothache)$ and $\mathbf{P}(Cavity|\neg toothache)$. [5]

(b) (i) Write down the Bayes rule. [2]

(ii) If $\mathbf{x}$ is the feature vector and ω is the class label, apply Bayes rule and write down the equation for $p(\omega|\mathbf{x})$. Explain qualitatively what this equation means, and how it can be applied for pattern classification. [4]

(iii) If each dimension (attribute) of $\mathbf{x}$ is independent of each other, write down the corresponding equation for $p(\omega|\mathbf{x})$. [3]

(iv) If each attribute is Normally distributed, how many parameters are there in total? What are they? [3]

(v) What is the name of this classifier? [1]

6. (a) What is the problem of a conventional Neural Network without hidden layer. (i.e. a single layer neural network, not counting the input layer). Explain your answer. [3]

(b) In a Convolutional Neural Network, 3 kinds of layers are commonly used. What are they? Explain their purposes. [6]

(c) What is the advantage of a Convolutional Neural Network over a Conventional Neural Network with the same number of layers? [2]

(d) Why the use of GPU (*Graphic Processing Unit*) is usually preferred in training a Convolutional Neural Network? [2]

END OF PAPER