THE UNIVERSITY OF HONG KONG

FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE

COMP7404 Computational Intelligence and Machine Learning

Date: May 13, 2017 **Time: 2:30pm - 4:30pm**

Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is the candidates' 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.

Calculator Brand: ____________________ *Calculator Model:* ____________________

Write your University No. on every page.

Answer all six questions in the space provided.

Special Note: Candidates are permitted to bring to the examination ONE sheet of A4-sized paper with printed or written notes on both sides.

Question No.	Weight (in %)	Your Mark (in %)
1	30	
2	30	
3	12	
4	10	
5	8	
6	10	
Total	100	

1. True / False Questions

For each True / False question, write down True or False. A correct answer is worth 2 marks, a missing answer is worth 0 marks and a wrong answer is worth -1 marks. You may provide an explanation, but this is optional.

(a) [*True* or *False*] The heuristic $h(n) = 0$ is consistent for every search problem.

(b) [*True* or *False*] An admissible heuristic never overestimates the cost to the goal.

(c) [*True* or *False*] When using alpha-beta pruning, it is possible to get an incorrect value at the root node by choosing a bad ordering when expanding children.

(d) [*True* or *False*] The graph-search version of A* will be optimal if an admissible heuristic function is used.

(e) [*True* or *False*] The greedy search algorithm is complete.

(f) [*True* or *False*] In local search there can be more than one global optimum.

(g) [*True* or *False*] The most-constrained variable heuristic provides a way to select the next variable to assign in a backtracking search for solving a CSP.

(h) [*True* or *False*] Depth-first search with a depth limit can be made to return the same solution as breadth-first search by iteratively increasing the depth limit.

(i) [*True* or *False*] The starting position for the game tic-tac-toe has a minimax value of 0.

(j) [*True* or *False*] Assume $A \succ B$, $D \succ A$, $B \succ L$, where $L = [0.5, C; 0.5, D]$. With rational preferences it follows that $B \succ C$.

(k) [*True* or *False*] If the policy has converged in value iteration, the values must have converged as well.

(l) [*True* or *False*] If the values have converged in value iteration, the policy must have converged as well.

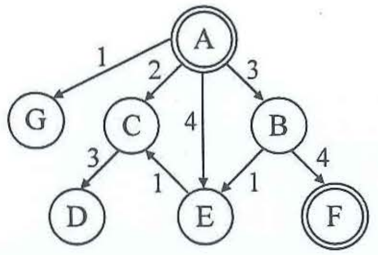
(m) [*True* or *False*] Value iteration is guaranteed to converge if the discount factor (γ) satisfies $0 < \gamma \leq 1$.

(n) [*True* or *False*] Q-learning can learn the optimal Q-function without ever executing the optimal policy.

(o) [*True* or *False*] For an infinite horizon MDP with finite number of states and actions with a discount factor γ, with $0 < \gamma < 1$, policy iteration is guaranteed to converge.

2. Search

(a) Consider the following search graph. A is the start state and F is the goal state. Assume that children are added to the frontier in alphabetical order.



i. Write down the order in which states will be explored by depth-first graph search. *(6 marks)*

ii. Write down the order in which states will be explored by breadth-first tree search. *(6 marks)*

iii. Write down the order in which states will be explored by uniform-cost tree search. If there is a tie use *first-in first-out.* *(6 marks)*

(b) The euclidean distance is an example of an admissible and consistent heuristic for the shortest path problem. Construct an admissible but inconsistent heuristic for some search problem. You may select any search problem. *(12 marks)*

3. Adversarial Search

Two players, player $\triangle$ and player $\triangledown$, find themselves in the following start state of an adversarial game.

$\triangle$			$\triangledown$
1	2	3	4

Player $\triangle$ moves first. The two players take turns moving, and each player must move to an open adjacent square in either direction. If the opponent occupies an adjacent space, then the player will jump over the opponent to the next open space, if any. For example, if player $\triangle$ is on square 3 and player $\triangledown$ on square 2, then player $\triangle$ may move back to square 1. The game ends when one player reaches the opposite end of the board. If player $\triangle$ reaches square 4 first, then the value of the game is +1; if player $\triangledown$ reaches square 1 first, then the value of the game is -1.

(a) Draw the minimax game tree and use the proper triangles to distinguish a *min* from a *max* node. Put terminal states in square boxes and loop states in double square boxes. Loop states are those that already appear on the path to the root. Write down the value of every state, you may write "*unknown*" for loop states. Can you indicate the optimal decision at the root of the tree? Explain. *(10 marks)*

(b) Consider an n-square game with $n > 2$ squares instead of the above game with 4 squares. How would the value at the root of the tree change? *(2 marks)*

4. Markov Decision Processes

Consider the following Grid World introduced in our lecture where an agent may move *north*, *south*, *east* or *west* and exit states are $a4$ and $b4$.

-1.11	-1.10	-0.94	+1	a
-1.11		-1.10	−1	b
-1.11	-1.11	-1.11	-1.10	c
1	2	3	4	

Actions are unreliable: 80% of time each action achieves the desired effect and 10% of the time the action moves the agent at right angles to the intended directions. If the agent runs into a wall, it stays in the same square. After running value iteration we have obtained the values shown in the figure. Note that values have been rounded to two decimal places. Let the discount factor $\gamma = 0.1$ and the reward R(s) = -1. Consider the two states $a3$ and $c4$.

(a) Calculate the Q-values of the two states. *(8 marks)*

(b) Perform policy extraction for the two states. *(2 marks)*

5. Reinforcement Learning

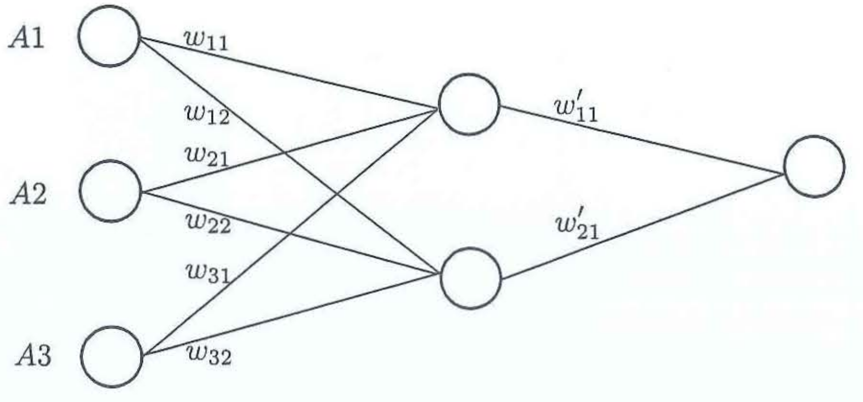
For Q-Learning to converge to the optimal Q-values ... *(8 marks)*

- ☐ It is necessary that the discount γ is less than 0.5.
- ☐ It is necessary that every state-action pair is visited infinitely often.
- ☐ It is necessary that the learning rate α (the weight given to new samples) is decreased to 0 over time.
- ☐ It is necessary that actions get chosen according to $\arg\max_a Q(s, a)$.

(Mark all that apply)

6. **Neural Networks**

Consider the problem of predicting the performance of a student in the final exam based on that student's performance in assignment 1 (A1), assignment 2 (A2) and assignment 3 (A3). We are going to use the following neural network.

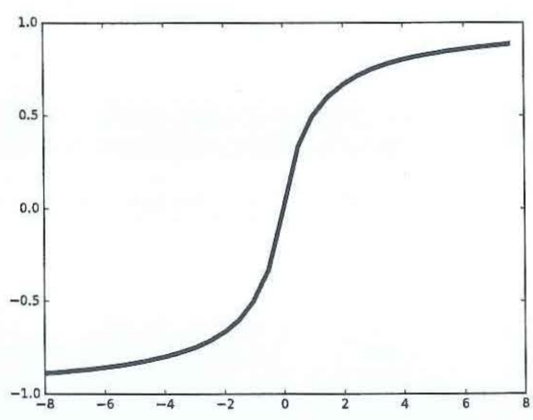


(a) Label the input nodes, neurons and synapses in the neural network above. *(2 marks)*

(b) Assume that training has been completed and the weight matrices are given as

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix} = \begin{bmatrix} 1.0 & 0.9 \\ 0.8 & 0.7 \\ 0.6 & 0.5 \end{bmatrix}, \mathbf{W}' = \begin{bmatrix} w'_{11} \\ w'_{21} \end{bmatrix} = \begin{bmatrix} 1.0 \\ 0.9 \end{bmatrix}.$$

We are going to use the softsign function, defined as $f(z) = \frac{z}{1+|z|}$, for the activation function of the neural network. It has the following shape.



Predict the final exam performance of a student that scored 0.6 (i.e., 60%) in A1, 0.5 (i.e., 50%) in A2 and 0.6 (i.e., 60%) in A3. Note that all input and output values of the neural network should be in $[0, 1]$. Show all your computations. *(8 marks)*

END OF PAPER

Please be reminded to write your university number on every page.

This page is intentionally left blank.
You may use it to draft your answers.