CS4725: Introduction to Artificial Intelligence/
CS6705: Foundations of Artificial Intelligence

Sample Final Exam

Name: ______________________________

Student ID: ______________________________

Signature: ______________________________

- Answer all questions in the space provided. Use the back of the page if necessary, but indicate clearly if anything on the back is to be marked.
- No calculators, books, notes or other aids are allowed.
- This is a 3-hour exam. You must stay for at least 30 minutes. Manage your time wisely.
- Read through the whole exam before starting, and ask for clarification if any questions are unclear. If you are still unsure of what a question is asking, state clearly what assumptions you are making before providing your answer.
- Have a great summer!

1.	/ 10
2.	/ 3
3.	/ 3
4.	/ 6
5.	/ 4
6.	/ 4
7.	/ 4
8.	/ 6
9.	/ 8
10.	/ 7
Total	/ 55

1. **(10 marks)** Short-answer / multiple-choice questions:

   (a) (2 marks) Near the beginning of the course, we talked about the idea of specifying an intelligent agent's task by providing a **PEAS** description. What did the letters **P**, **E**, **A** and **S** stand for?

   (b) (1 mark) Suppose you want to represent some task as a **search problem**, as you did on Assignment 1 for the *Missionaries and cannibals* problem. Your specification must include six components: the **set of states**, the **initial state**, the set of **actions** available, the **transition model**, the **path cost function** and which other component?

   (c) (1 mark) **Optimality** and **completeness** are two of the four criteria that we used to evaluate uninformed search algorithms. What are the two other criteria?

   (d) (2 marks) Suppose we have a search space in which every action has a cost of 1 and in which the branching factor is $b = 4$. Sequences of actions exist that are up to length 20, but we know in advance that all solutions will occur at depth $d = 10$. Which of the following algorithms would be both complete and optimal? Circle all that apply.

      i. breadth-first search
      ii. uniform-cost search
      iii. standard depth-first search
      iv. depth-limited search, $l = 8$
      v. depth-limited search, $l = 10$
      vi. depth-limited search, $l = 12$
      vii. iterative deepening search

(e) (2 marks) Suppose that you are solving some search problem in which all action costs are positive. You have two heuristic functions, h_1 and h_2, and you know that both of them are admissible. You can assume that $h_1(n) \geq 0$ and $h_2(n) \geq 0$ for all states n.

From the information provided above, it follows that all of the heuristics below are also admissible. Based on our discussion in class, which one of the heuristics below would lead to the most efficient performance when used as the heuristic in A* search? Circle one answer only.

i. $h_3(n) = 0$ for all states n
ii. $h_4(n) = 0.5(h_1(n) + h_2(n))$ for all states n
iii. $h_5(n) = \min\{h_1(n), h_2(n)\}$ for all states n
iv. $h_6(n) = \max\{h_1(n), h_2(n)\}$ for all states n
v. $h_7(n) = 0.75h_1(n)$ for all states n

(f) **(2 marks)** Suppose that we have a **neural network unit** that has two boolean inputs, x_1 and x_2, and a single boolean output, y.

Like the units we designed in class for the boolean AND and OR functions, this neural network unit also uses a fixed input, x_0, set to 1. It uses a threshold activation function such that it outputs 0 if the weighted sum is negative and outputs 1 otherwise.

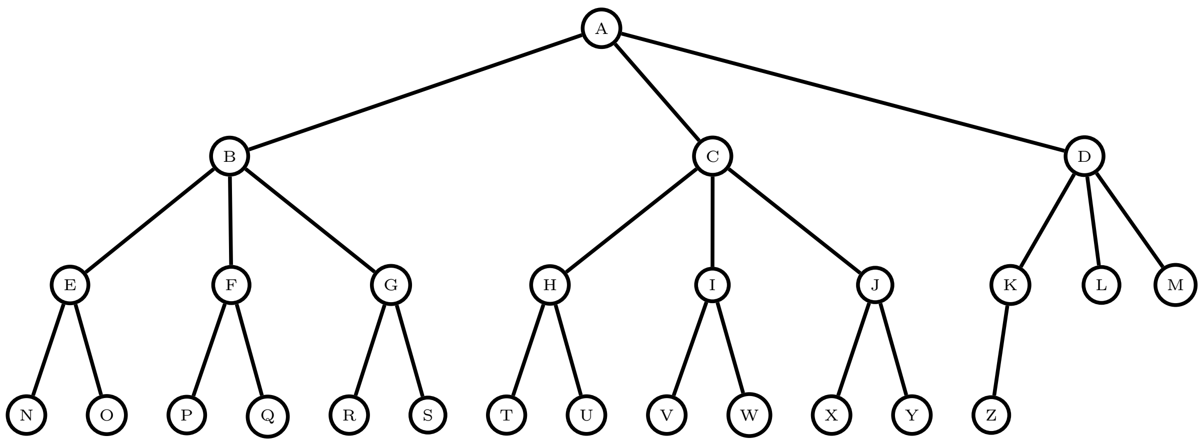
The neural network algorithm is trying to learn the function $x_1 \Rightarrow x_2$, shown below.

x_1	x_2	$x_1 \Rightarrow x_2$
0	0	1
0	1	1
1	0	0
1	1	1

Which of the following sets of learned weights would produce correct behaviour for this neural network unit?

i. $w_0 = -3$, $w_1 = 2$, $w_2 = 2$
ii. $w_0 = 1$, $w_1 = -2$, $w_2 = 2$
iii. $w_0 = -2$, $w_1 = 3$, $w_2 = 3$
iv. $w_0 = 1$, $w_1 = 2$, $w_2 = -2$

2. **(3 marks)** Consider the following search tree, where node J is a goal node. List the order in which nodes would be selected for expansion by each of the following search algorithms. (Note: Even if a node has no children, it should be included on your list when it would be selected by the algorithm to check if it has children.)

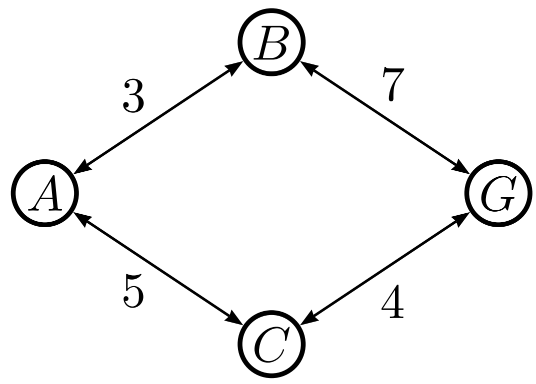


- Depth-first search:

- Depth-limited search, $l = 2$:

- Iterative deepening search:

3. **(3 marks)**



Consider the very simple state graph shown above, where G is the goal state. Each edge can be traversed in both directions and is labelled by its cost. You have been given the heuristic function h below, but the value of $h(A)$ is unknown. You can assume that it is a non-negative value.

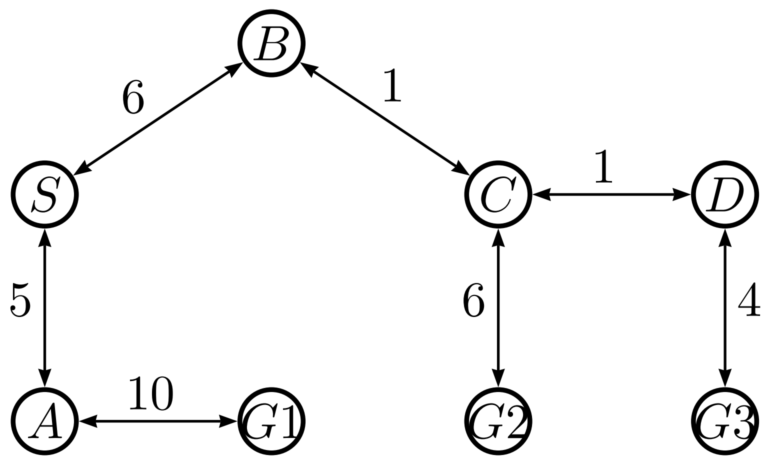
Node n	A	B	C	G
$h(n)$	?	5	2	0

(a) What range of values for $h(A)$ would make h an admissible heuristic?

(b) What range of values for $h(A)$ would make h a consistent heuristic?

4. **A* search (6 marks)**

Consider the problem of getting from node S to a goal node ($G1$, $G2$ or $G3$) in the graph below.



Trace the operation of **A* search** with a heuristic function that has values $h(S) = 8$, $h(A) = 1$, $h(B) = 6$, $h(C) = 1$, $h(D) = 2$, $h(G1) = 0$, $h(G2) = 0$, $h(G3) = 0$.

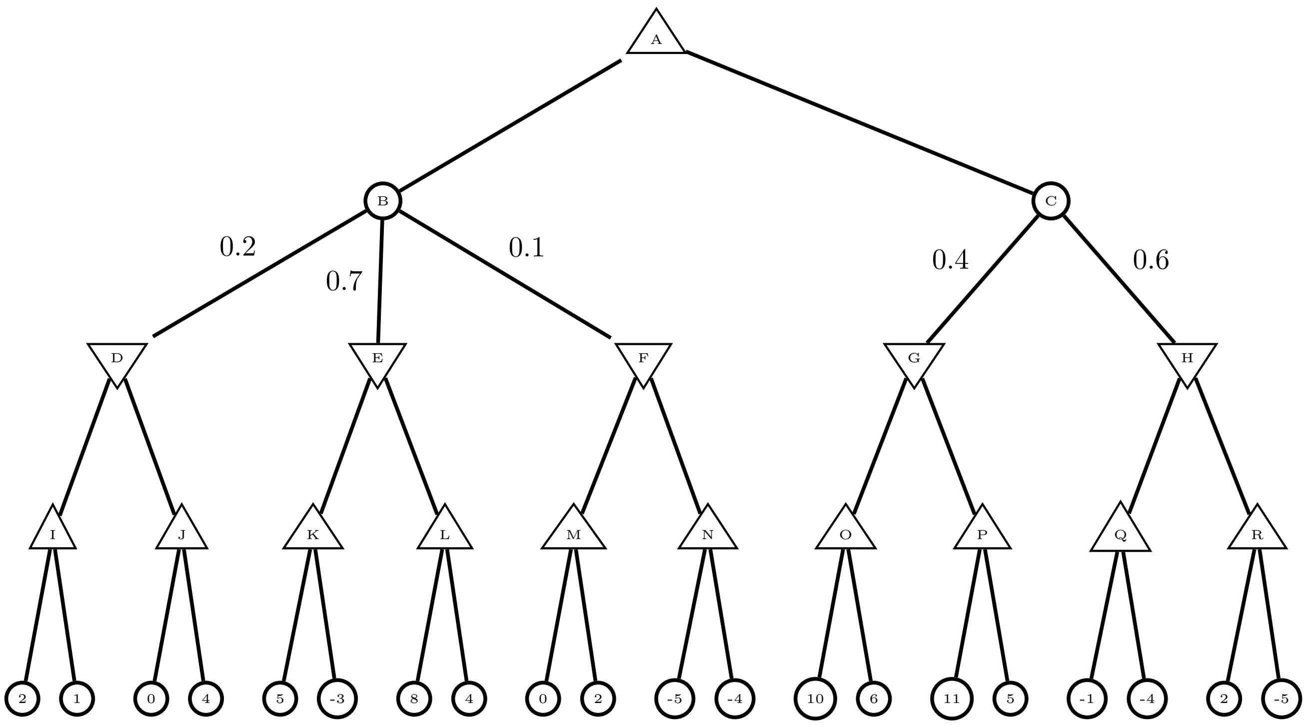
Draw the search tree that would be built; beside each node in the search tree, show the calculation of its f-value. Also, fill in the table below. (If two nodes on the frontier have the same f-value, then break ties in favour of the node with the lower h-value.)

Node selected for expansion	Priority queue (show nodes and their f-values)

Solution path found:

Cost of solution path:

5. **(4 marks)** Consider the game tree below, representing a game in which there is an element of chance involved. Nodes A and I-R are MAX nodes, and nodes D-H are MIN nodes. Nodes B and C are chance nodes, with the probabilities of different chance events indicated on the edges leaving the nodes.

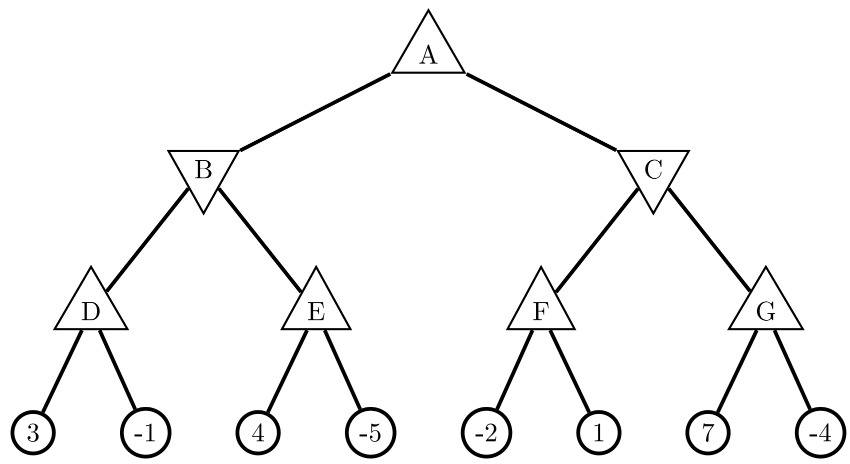


Write the **expectiminimax value** of each node beside the node. Show your calculations for the chance nodes here:

What is the best move for MAX at node A: to go to node B or C?

6. **(4 marks)** In the game tree below, circle all nodes or subtrees that will not have to be visited during a depth-first search using the minimax algorithm with **alpha-beta pruning**. You do not have to provide any explanation.

   A, D, E, F and G are MAX nodes. B and C are MIN nodes.



7. **(4 marks)** Consider the following full joint distribution, similar to the example given in Chapter 13 of our course textbook.

   For example, this table shows that $P(\neg\text{toothache}, \neg\text{cavity}, \neg\text{catch}) = .58$.

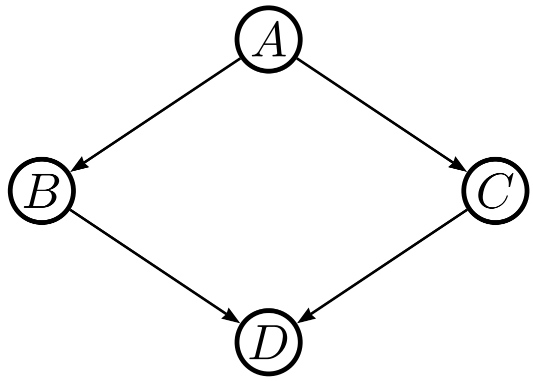
	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	.11	.01	.07	.01
¬cavity	.02	.06	.14	.58

   Use the values in the full joint distribution above to calculate the following:

   (a) $P(\text{cavity}) =$

   (b) $P(\text{cavity} \mid \text{toothache}) =$

8. **(6 marks)** Consider the following Bayesian network, where A, B, C and D are all boolean variables (true or false).



If you are using the method of **exact inference by enumeration** (recall the burglary alarm example from class), write the formulas that you would need to use in order to calculate $P(b \mid d, \neg a)$ and $P(\neg b \mid d, \neg a)$. You do not have to do any calculations (in fact, you are not even provided with the probability values that you would need in order to perform the calculations); you must just show the formulas for the calculations that would have to be done.

$P(b \mid d, \neg a) \ =$

$P(\neg b \mid d, \neg a) \ =$

9. **Game Trees and Decision Theory (8 marks)**

[Credit to Scott Buffett] Imagine that you are a criminal who has just been caught. While discussing a possible plea bargain, the crown prosecutor invites you to play "Squeal or No Squeal," a new fun way to settle criminal cases.

First, the prosecutor gives you a suitcase containing either a black ball or a white ball. Before knowing what is in the suitcase, you must choose either "Squeal" and confess your crime and name others involved, or "No Squeal" and keep quiet. Then you open your suitcase. If you chose "Squeal", you get 9 years in prison if the ball is black and 4 years in prison if the ball is white. If you chose "No Squeal", you get 10 years in prison if the ball is black, but if the ball is white, you get to play "Squeal or No Squeal" one more time with a new suitcase. This second time, "Squeal" will get you 3 years for sure, and "No Squeal" will get you 4 years if the ball is black or 0 years if the ball is white.

Consider that your utility for y years in prison is $1 - 0.1y$. Assume the probability of getting each colour of ball is 0.5.

(a) (4 marks) Draw a tree to represent this problem and show your work in using the tree to determine your best strategy and its expected utility.

(b) (4 marks) Now, suppose that you can find out before the game begins what colour of ball is inside the first suitcase **or** what colour of ball is in the second suitcase. Which piece of information should you choose? Draw a modified tree for each of these two situations and then show your work in calculating the value of information in each case.

10. **(7 marks)** Suppose you have been asked to use **decision tree learning** to generate a decision tree that will correctly categorize the following training examples. Attributes A, B and C are the input attributes, and Goal is a single Boolean output variable.

    Note that there are 4 positive (Goal = 1) and 4 negative (Goal = 0) training examples.

A	B	C	Goal
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

    (a) What is our initial **entropy** (the uncertainty of the random variable Goal)?

    (b) As we begin to build our decision tree, we must decide which of attributes A, B and C would be the best splitting attribute to use at the root of the tree.

    Write a mathematical expression showing the **information gain** associated with splitting on attribute C.

    You can leave values like $\log_2(\frac{1}{3})$ in your expression without computing these values. However, your answer should show all of the quantities that would have to be added, subtracted, multiplied and divided in order to compute the information gain associated with splitting the training set on attribute C.

    (c) Without computing the actual values for the information gain associated with splitting on attributes A, B and C, explain briefly why A and B would be poor choices for a splitting attribute at the root of the decision tree.