CMSC 471 Final

Name:

1. **Short Answer**

(4) (a) What goal does the ID3 algorithm attempt to accomplish when making decision trees? Is ID3 always guaranteed to find the optimal decision tree that is consistent with a given training set?

(4) (b) In state-space search, one can perform the goal test when a state is generated, or when it is expanded. Which of these should you do to guarantee the optimality of A* search, and why?

(3) (c) What is the difference between supervised and unsupervised learning?

(5) (d) A machine learning algorithm has 75% accuracy on the training set used by researchers. Why is that not enough to evaluate the algorithm? When might 75% not be a particularly notable accuracy? What other information might you want when evaluating the merit of their process?

(3) (e) What factors influence the likelihood that k-means will converge to clusters that actually capture the structure of the data (you don't need to use all the space)?

(5) 2. **Utility Functions**

In general, one can think of each player in a two-player game as having a utility function. Let's call player A's utility function, which maps states into real-valued utilities, $U_A : S \to \mathcal{R}$. Player B's utility function is $U_B : S \to \mathcal{R}$.

In a zero-sum game, which includes all of the games we have looked at so far,
$\forall s \in S, U_A(s) = -U_B(s)$; that is, the players always have exactly opposite utilities.

Suppose that the players in a two-player game are not playing a zero-sum game. (That is, U_A and U_B are not opposites.) How could the minimax algorithm be modified to handle this case?

3. **Machine Learning**

You have a shiny new iPad, a shiny new developers license from Apple, and a friend with an idea for a great game that you think will be a big seller. You know that making a great game requires more than just shiny graphicsit requires good AI! However, your friend has left the country for the summer and forgot to write down the formula for scoring the game. All you have is a set of labeled examples, given below.

Example	**Foos**	**Blargs**	**Goozles**	**Sparkles**	**Zazzles**	**Outcome**
1	1	2	1	2	3	Win
2	1	3	1	1	2	Tie
3	3	2	2	3	4	Win
4	4	0	1	1	2	Win
5	1	4	0	2	4	Loss
6	0	2	0	0	1	Loss
7	3	1	1	1	2	Win
8	2	4	3	2	1	Win
9	1	1	2	1	1	Win
10	1	3	1	1	2	Tie
11	2	4	0	1	4	??

(5) (a) Suppose a decision tree is built with a root node of "Is the number of Foos ≥ 2?" Suppose the agent is deciding whether to use "Are the Blargs ≥ 3?" or "Are the Goozles ≥ 1?" in the next split that it needs to construct in recursively building the tree. Draw these two alternative trees, showing which examples end up at each node (root, interior, and leaf nodes), and how many instances of each class appear at each node. Use ONLY examples 1–10.

(5) (b) Compute the information gain (at the second-level split) for each of the two attributes under consideration in the previous question. Again, use ONLY examples 1–10. Which attribute should the decision tree algorithm choose, if it is using information gain as the heuristic?

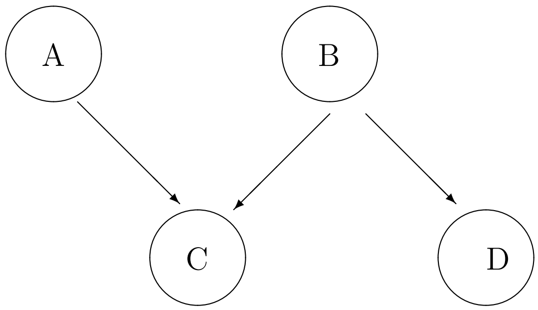
(3) (c) What does k-nearest neighbor predict for Example #11 (the unlabeled instance) using $k = 3$, if similarity is computed using *only* the attribute **Zazzles**? Justify your answer.

4. **Bayesian Networks**

Consider the following Bayesian network:

$P(a) = 0.7$

$P(b) = 0.4$



$P(d|b) = 0.2$
$P(d|\neg b) = 0.5$

$P(c|a, b) = 0.9$
$P(c|\neg a, b) = 0.6$
$P(c|a, \neg b) = 0.3$
$P(c|\neg a, \neg b) = 0.1$

A, B, C, and D are all Boolean variables. In the probability expressions below, the appearance of a variable name in lower-case (e.g., a) means that that variable takes on the truth value `True`. The lower-case negation of the variable (e.g., $\neg a$) means that that variable takes on the truth value `False`.

(3) (a) **BN Inference**

Compute $P(a \wedge \neg b \wedge c \wedge \neg d)$ from the Bayesian network. Show your work.

(5) (b) **Enumeration**

Write the summation to compute $P(c)$ directly the joint probability distribution. (You do *not* need to give a numeric probability.)

How many lookups into the conditional probability tables does the computation require, assuming no caching, and assuming every parameter (not just independent parameters) are explicitly retrieved from the CPTs? Explain your answer.

(4) (c) **BN Structure**

If an edge were added from C to D, would the resulting graph still be a Bayesian network? If so, how many additional independent parameters would the new BN require? If not, why isnt it a BN?

5. Consider the game:

19, 39	44, 13	16, 8	64, 57
47, 59	75, 25	64, 88	74, 63
13, 45	102, 28	81, 84	93, 13
36, 43	83, 17	24, 77	83, 30

(1) (a) What is the move that maximizes social welfare?

(3) (b) Indicate the nash equilibriums, if any.

(3) (c) List all pareto optimal states.

(3) (d) What would be the minimum alteration to the matrix (smallest magnitude change) that would create a dominant strategy for the player 2? (Player 2 is the one who chooses the columns, their score value is listed second.)

6. **MDPs**

Consider this simple grid world:

S		
	-1	
		5

An agent starts in the upper left ($[1, 1]$), and has two actions it can take in each state: Right or Down. Right has the following effects:

- With probability .9, move one step to the right (or stay in the same location if the agent is in the rightmost column).
- With probability .1, move one step down (or stay in the same location if the agent is in the bottom row).

Down always moves down (with probability 1), unless the agent is in the bottom row, in which case it stays in the same place. Both actions always have reward -1, regardless of the resulting state. Locations $[2, 2]$ (the center square) and $[3, 3]$ (the lower right square) are terminal states, with utility (value) as shown. The agent uses a discount factor $\gamma = 0.9$.

(6) (a) Value Iteration Using a random policy (i.e., the probability of each action is 0.5), and starting with $V(s) = 0$ for all non-terminal states, the agent applies value iteration to compute the values associated with each state. After two iterations, it has computed the following state values ($V(s)$):

-2.3	-1.9	-.68
-1.9	-1	2.1
-1.1	-1.5	5

where the terminal states (which do not change during value iteration) are shown as shaded squares. Use value iteration (just a single round) to compute the *next estimate* for state $[1, 1]$ (that is, the upper left square). *Show your work.*

(4) (b) Policy Estimation

Given the estimated values at the beginning of the previous question (not your updated estimate), what is the current best policy for the agent to adopt? Show your answer as arrows on this diagram:

-2.3	-1.9	-.68
-1.9	-1	2.1
-1.1	-1.5	5

7. **Resolution Theorem Proving**

(10) (a) **First-Order Predicate Logic (8 pts)**

Convert the following sentences into first-order predicate logic:

1. HAL is intelligent.
2. HAL is not human.
3. Intelligent beings have intentionality.
4. Beings with intentionality are conscious.

Use only the predicates intelligent(x), human(x), intentional(x), and conscious(x), and the constant HAL. (You may abbreviate the predicates as "intel," "h," "intent," and "c" if you wish). All variables can be assumed to have the domain of "all beings."

(7) (b) **Conjunctive Normal Form**

Convert the KB (i.e., the logical sentences from the previous question) into conjunctive normal form, numbering them as in the list above. (If any sentence above corresponds to more than one sentence in the KB, number them (for example) 2a, 2b, etc.)

(3) (c) **Goal Negation**

Write the goal, "There is a conscious non-human," in first-order logic, then write the negated goal in conjunctive normal form.

(10) (d) **Resolution Proof**

Prove using resolution on the KB and negated goal from the previous questions that there is a conscious non-human being. You may use either a tree form or a sequential form for your proof, but you should be sure to show all variable bindings.

If you can't find a resolution tree, I will give partial credit if you show a series of valid logical inferences (and justifications, i.e., inference rules) to prove the goal from the KB. You may indicate sentences from the KB by number rather than copying them out here.