CS472
Foundations of Artificial Intelligence

Final Exam

December 19, 2003
12-2:30pm

Name:

(Q exam takers should write their Number instead!!!)

Instructions: You have 2.5 hours to complete this exam. The exam is a closed-book, closed-notes exam.

#	description	score	max score
1	Search	____	23
2	Planning	----	22
3	Machine learning I	____	16
4	Game playing	____	21
5	Machine learning II	----	20
6	Logic and Knowledge-based systems	____	15
7	Natural language processing	----	8
	Total score:	________ --------	/ 125

Search (and a bit o' machine learning): (23 points total)

1. (3 pts) True or False: Best-first search can be thought of as a special case of A*. Briefly explain your answer.

2. (3 pts) True or False: Breadth-first search can be thought of as a special case of uniform-cost search. Briefly explain your answer.

3. (3 pts) Under what conditions does A* search produce the optimal solution?

4. The ID3 algorithm for top-down induction of decision trees can be viewed as a greedy, heuristic search for the "best" decision tree. Describe how to formulate ID3 as such a search problem below:

   - (3 pts) What does each state represent?

   - (2 pts) What is the initial state?

- (3 pts) Describe the goal state/test.

- (4 pts) Explain how the successor function would operate.

- (2 pts) What is the heuristic function?

Planning (and knowledge representation): (22 points total)

In this question we'll consider the perennial Christmas Candy problem, in which there is a child in a room with some candy that is hanging on the Christmas tree, out of reach of the child. A stool is available that will enable the child to reach the candy if he climbs on it. Initially, the child is at location A; the candy at B, and the stool at C. The child and stool have height *low*, but if the child climbs onto the stool he will have height *high*, which is the same height as the candy. The actions available to the child include *GO*ing from one location to another; *PUSH*ing an object from one place to another, *CLIMB*ing onto an object, *GRASP*ing an object. Grasping results in holding the object if the child and object are in the same place at the same height. The goal, of course, is for the child to be holding the candy. (Note: for the problems below, there is no need to explicitly represent the lovely Christmas tree.)

1. (2 pts) Write down the initial state description in first-order logic.

2. (2 pts) Now specify the initial state description using situation calculus.

3. (3 pts) Using situation calculus, write down the **possibility** and **effect** axioms for the *GRASP* action.

4. (3 pts) Show the *GO* action schema using the STRIPS language.

5. (3 pts) Show the initial partial-order plan (either pictorially or as a plan data structure) for the Christmas Candy problem.

6. (3 pts) Show one possible non-initial partial-order plan for the Child and Christmas Candy problem.

7. (3 pts) True or False: The STRIPS representation language for planning solves the *representational frame problem*. Provide an explanation for your answer.

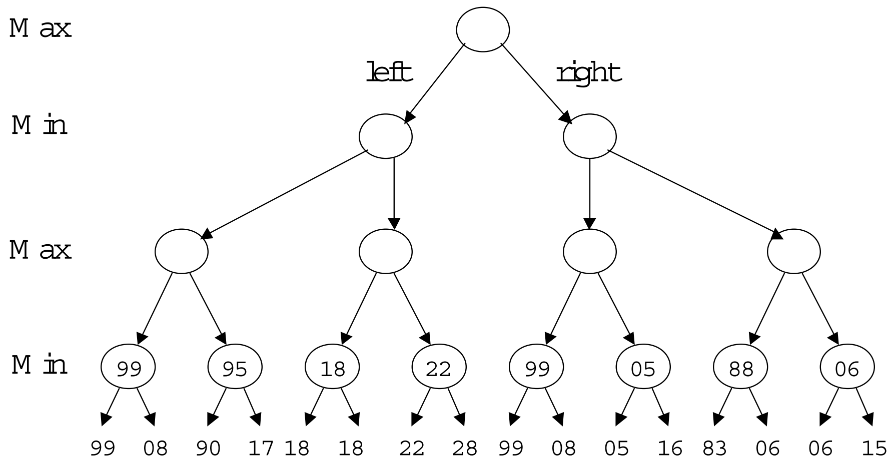
Machine learning I: (16 points total)

1. (3 pts) Consider the class of neural networks with inputs $x_1 \ldots x_n$, which are either 0 or 1. Specify the smallest network that computes the **majority function** for 5 input nodes (i.e. the network should output a 1 if at least half the inputs are 1's and should output a 0 otherwise). Networks should specify the weights on all links and the activation function g at each node.

2. (3 pts) Draw a decision tree that represents the **2-of-3 majority function**. Again, assume binary input values.

3. (5 pts) Show how any propositional Boolean formula representable as the conjunction of p positive literals and n negative literals can be represented by a perceptron. Assume that the perceptron uses thresholded units.

4. (5 pts) Describe *why* and *how* you might apply simulated annealing to train a neural network.



Game playing: (21 points total)

In the game tree above, the numbers indicate values for the heuristic evaluation function.

1. (3 pts) Should Max choose the **left** or **right** branch *if the depth of the search is 3* (i.e., Max chooses, Min chooses, Max chooses)? Show the minimax values in the tree.

2. (3 pts) Should Max choose the **left** or **right** branch *if the depth of the search is 4*? Show (and underline) the minimax values in the tree.

3. (3 pts) Explain why the minimax values might change when computed via a depth-3 vs. depth-4 search.

4. (3 pts) Using the full tree, show one example of alpha-beta pruning by circling the pruned nodes. Assume that nodes are evaluated from left to right.

5. (3 pts) Does alpha-beta pruning always select the same line of play as standard minimax? Briefly explain your answer.

6. (3 pts) Explain the role of **chance nodes** in game tree search.

7. (3 pts) Provide an equation that shows the “expectiminimax” value computation for chance nodes.

Machine learning II: (20 points total)

3	?	-0.8	**-1**
2	0.5	**+1**	-0.3
1	0.1	0.2	-0.3
	1	2	3

1. **Reinforcement learning**. Consider the 3x3 world in the figure above in which the agent receives a reward of -1 upon reaching state (3,3) and a reward of +1 upon reaching state (2,2). The reward in all other states is 0. In addition, the figure shows the current utility values for each state in the environment. The "?" in state (1,3) indicates that the utility value of this state is currently unknown because the state has not yet been visited.

   (a) (5 pts) Calculate the new utility value for state (1,3) (i.e. the state with the "?") using the *Widrow-Hoff naive updating algorithm* (also called the direct utility estimation method) after the agent executes the following trial: (1,1) → (1,2) → (1,3) → (2,3) → (2,2).

   (b) (5 pts) Starting again with the utility values in the figure, use the *temporal difference learning rule* (with the learning rate α set to 0.1) to calculate the new utility value for state (1,3) (i.e. the state with the "?") after the agent executes the same trial: (1,1) → (1,2) → (1,3) → (2,3) → (2,2). (If you are using a version of TD learning that has a discount factor, set it to 1.)

2. (10 pts) **Genetic algorithms.** Consider the following *fitness function*:

fitness (bitstring) = number of 1's in the bitstring where **both** *adjacent bits are 0's.*

For example, fitness ("010110100") = 2; fitness ("100011011") = 0; and fitness ("010101010") = 4. (Notice that 1's in the first or last position in the string are not counted in the fitness function, even if adjacent to a 0.) Assume the design of our genetic algorithm is:

(a) Create an initial population containing 4 random 9-bit strings.

(b) Discard the 2 least-fit ones (break ties randomly).

(c) Do a single-point cross-over using the 2 most fit. The 2 children that result and their parents constitute the next generation.

(d) Randomly mutate 1 bit in 1 string in the population.

(e) Go to step (b).

Start with the initial population below and show what the next two (2) generations might look like. Explain your reasoning.

Generation 0

0	1	1	1	1	0	1	1	0
0	1	1	0	0	1	0	1	1
1	0	1	1	0	1	1	1	0
0	0	0	0	1	0	1	0	1

Generation 1 **Explanation**

Generation 2 **Explanation**

Logic and Knowledge-based Systems (15 points total)

1. (3 pts) Translate the following logical statement into *good, natural* English (no x's and y's):

   $$\forall x, y, l \;\; SpeaksLanguage(x, l) \wedge SpeaksLanguage(y, l) \rightarrow Understands(x, y) \wedge Understands(y, x)$$

2. Translate into first-order logic the following sentences. (Remember to define all predicates, functions, and constants. For full credit, avoid using any *ReallyLongPredicateNames*.)

   (a) (3 pts) If someone understands someone, then she is that someone's friend.

   (b) (3 pts) Friendship is transitive. (Hint: Do not use the predicate *transitive* here.)

3. (6 pts) Suppose that Ann and Bob speak French and Bob and Cal speak German. Prove *using resolution* that Ann is Cal's friend, using as axioms the sentences from parts 1 and 2. (Be sure to convert the knowledge base to CNF; show any unifications required. You may abbreviate any symbols as necessary.)

Natural Language Processing (8 points total)

A major problem in the design of natural language understanding systems is dealing with the massive amount of ambiguity at each phase of the communication process. For each disambiguation problem listed below, name the primary stage of analysis (i.e. intention, perception, synthesis, hearer disambiguation, hearer incorporation, semantic interpretation, parsing, pragmatic interpretation) in which the problem would be encountered:

1. (2 pts) Determining whether the utterance was "It's hard to recognize speech" or the very similar sounding "It's hard to wreck a nice beach".

2. (2 pts) Determining if *bank* refers to a river bank or a financial bank in: "I went to the bank on Monday."

3. (2 pts) Given the sentence "Jan saw the man with the binoculars", determining whether (1) it is the man who has the binoculars, or (2) Jan used binoculars to view the man.,

4. (2 pts) Determining whether or not "George" and "he" refer to the same real-world entity in the sentence: "Because George wanted to maintain his multi-million-dollar income, he decided not to divorce his wife."