FALL 2000
COMPUTER SCIENCES DEPARTMENT
UNIVERSITY OF WISCONSIN – MADISON
PH.D. QUALIFYING EXAMINATION

Artificial Intelligence

Monday, September 18, 2000
3:00 – 7:00 PM
1213 Engineering Hall

GENERAL INSTRUCTIONS:

1. Answer each question in a separate book.
2. Indicate on the cover of *each* book the area of the exam, your code number, and the question answered in that book. On *one* of your books, list the numbers of *all* the questions answered. *Do not write your name on any answer book.*
3. Return all answer books in the folder provided. Additional answer books are available if needed.

SPECIFIC INSTRUCTIONS:

Answer:

- either one (1) of the questions G540-1 or G540-2, *and*
- both (2) questions in the section labeled B760, B766, or B780 corresponding to your chosen focus area, *and*
- any two (2) additional question in the sections B731, B760, B766, and B780, where these two questions need *not* come from the same section, *and*
- both (2) questions in the section labeled A7xx that corresponds to your focus area.

Hence, you are to answer a total of *seven* (7) questions.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tries to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced that a problem has been stated incorrectly, mention this to the proctor. If necessary, the proctor can contact a representative of the area to resolve problems during the *first hour* of the exam. In any case, you should indicate your interpretation of the problem in your written answer. Your interpretation should be such that the problem is nontrivial.

Answer either one (1) of the questions G540-1 or G540-2.

G540 – Introduction to AI Questions

G540-1. Assume a search space where a solution is found at depth d, the branching factor is b, and the maximum depth is m.

(a) What is the worst-case time complexity and space complexity for (i) breadth-first, (ii) depth-first and (iii) iterative deepening searches of this space?

(b) If we assume that there are multiple solutions, do any of the three search strategies guarantee finding a minimal depth solution? Which one(s)?

(c) If we now assume that the maximum depth is unbounded, are any of the three search strategies complete? Which one(s)?

G540-2. Express the following in first order logic:

(a) Everyone likes ice cream.

(b) One's mother is one's female parent.

(c) Not all students take both History and Biology.

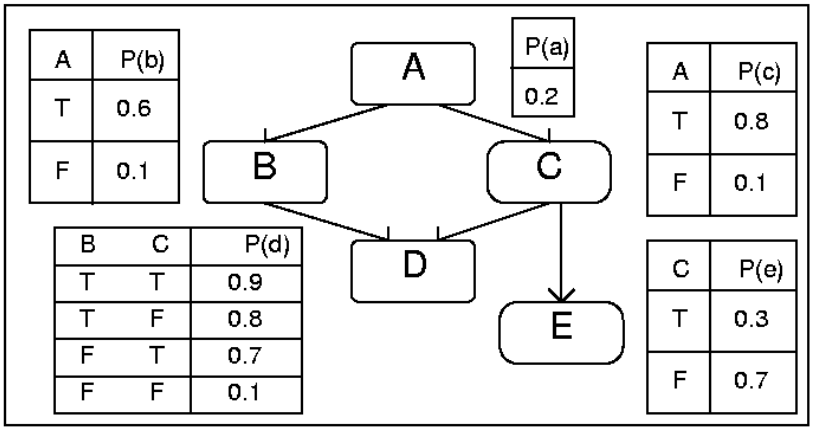
(d) There is exactly one apple.

(e) Selling something changes who owns it.

Answer both (2) of the questions in the section labeled B7xx that corresponds to your chosen focus area. Also answer any two (2) additional questions in any of the other sections (these two questions need NOT occur in the same section).

B731 – Basic Questions

B731-1. In the Bayesian Network below, the variables are A, B, C, D, and E, and they are all Boolean. In the conditional probability tables (CPTs), the notation P(a) denotes the probability that A is set to *True*, with similar meanings for P(b), P(c), etc. Use variable elimination to compute the probability distribution over A given the observations that D is *False* and E is *True*. You may eliminate the variables in any order you wish, but please show your work. (If you have neither a calculator nor time to perform the computations, then please simply write out in full the arithmetic expressions to be evaluated.)



B731-2. Write pseudocode for the GSAT algorithm for satisfiability (a paragraph description in English is acceptable if it is clearly written). What (if any) are the advantages of this algorithm over a greedy hill-climbing search? Show the first 3 steps of running GSAT on the CNF formula

$$(x_1 \vee \overline{x_2}) \wedge (x_2 \vee \overline{x_3}) \wedge (\overline{x_1} \vee \overline{x_3})$$

if GSAT happens to begin by randomly choosing the assignment mapping x_1 to 0, x_2 to 1, and x_3 to 1. If a step involves a random choice, please say this, state the alternatives, choose any one of them, and continue.

B760 – Machine Learning Basic Questions

B760-1. (a) Is it possible to *overfit* noise-free data? Show an example or argue that overfitting only occurs in the presence of noise.

(b) Discuss how one might use each of the following to reduce overfitting:

- tuning sets
- MDL principle
- ensembles

(c) For two (2) of the three approaches listed above, discuss (i) one strength relative to the other approaches and (ii) one weakness compared to not using any overfitting-avoidance measure.

B760-2. Would it be possible to use the Naive Bayes approach for Q-learning? Explain how this could best be done and describe the major strength and weakness of such an approach.

B766 – Computer Vision Basic Questions

B766-1. A commonly used 1D smoothing filter is the Gaussian:

$$g_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{x^2}{2\sigma^2})$$

(a) Repeated convolutions with a 1D Gaussian with a particular standard deviation, σ_G, is equivalent to a single convolution with what type of filter? Your answer should include the mathematical form of the equivalent filter, the approximate size of the filter, whether the filter is linear or not, and what the main purpose of this repeated convolution is.

(b) What property of the coefficients of a particular discrete approximation of a 1D Gaussian ensures that regions of uniform intensity are unaffected by smoothing?

(c) A Gaussian filter is often used to smooth an image as part of an edge detection procedure. What factors affect the choice of an appropriate value for σ?

B766-2. (a) Specify three matching constraints that can be used to find point correspondences in stereo computer vision.

(b) A pair of cameras is arranged symmetrically with the two optical axes coplanar and making an angle of 45°. Sketch the family of epipolar lines in the left and right images. Assume both cameras have the same focal length *f*.

(c) What is the result of computing (i) sum-of-squared-differences (SSD), and (ii) cross-correlation between an arbitrary pattern and a perfectly black pattern (i.e., all 0s) over an image window? Does this suggest a bias by one of these match methods that is not present with the other one?

B780 – Robot Motion Planning Basic Questions

B780-1. Consider a point mobile robot equipped with a compass which allows it to maintain a specific direction - say, towards the north. The robot's objective is to escape from a finite planar maze that consists of a finite number of disjoint obstacles, each of finite perimeter. The robot has a counter, which is turned on when the robot meets an obstacle and starts integrating the turning angles. Assume that all the obstacle boundaries are at right angles. Then, turning left corresponds to +90 degrees or adding 1 to the counter. Similarly, turning right corresponds to -90 degrees or subtracting 1 from the counter. The robot will have escaped from the maze when it reaches an infinite straight line that stretches in the east-west direction and is located outside and to the north of the maze. Consider the following algorithm:

Step 1. Walk straight toward the north until an obstacle is encountered.

Step 2. Turn left leaving the obstacle on your right.

Step 3. Follow the obstacle boundary until the value of the counter is zero. Go to step 1.

Does the algorithm converge? If your answer is "yes", prove it. If your answer is "no", give a counterexample.

B780-2. Consider a three-dimensional, two link RR arm manipulator (R = "revolute"). The axes of the two joints intersect and are perpendicular to each other. How many solutions to the inverse kinematics equations for the two joint angles are possible? What is the shape of this arm's work space and its configuration (image) space? Draw a sketch or explain.

Answer both (2) of the questions in the section labeled A7xx that corresponds to your chosen focus area.

A760 – Machine Learning Advanced Questions

A760-1. Sketch a decision tree induction algorithm designed for problems where comprehensiblity of learned solutions is important, but standard algorithms produce very large, hard-to-understand trees. The algorithm should allow the user to control the tree size by either (i) specifying the maximum number of internal nodes in the final tree, or (ii) interrupting the tree-growing process at any time to obtain the current tree. For any specified tree size, we would like this algorithm to return a tree that is reasonably accurate on the training data. Compare and contrast your algorithm to ID3. Why wouldn't it be a good idea to use ID3, but simply limit it's tree-growing process to the specified number of nodes (or allow it to be interrupted)?

A760-2. Consider a machine learning problem in which the unknown target concept is a directed graph, where each node has the color required to be "red," "black," or "don't care" (color of the node doesn't matter). Each example also is a directed graph, with every node colored either red or black. A target concept labels an example as "positive" if the target is a subgraph of the example (such that the colors match); otherwise, the target labels the example as "negative." For illustration, below is a target concept, a positive example of the concept, and a negative example of the concept. (The checkered nodes are "red" and the clear node in the target is a "don't care.") Show the background knowledge and data you would use to provide the two examples to FOIL. Describe how you would represent these examples for a decision-tree learner. Discuss one major strength and one major weakness of each of these two approaches.