SPRING 2001
COMPUTER SCIENCES DEPARTMENT
UNIVERSITY OF WISCONSIN – MADISON
PH.D. QUALIFYING EXAMINATION

Artificial Intelligence

Monday, February 5, 2001
3:00 - 7:00 PM
115 Psychology

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. Express the following English sentences using First-Order Predicate Calculus. You must use situation calculus for the last sentence.

(a) *Some red cookie weighs less than every green one*.

(b) *No Martian plays basketball*.

(c) *Each Picasso painting is described in a separate chapter of the text book for Art 101*.

(d) *Driving a car from Madison to Chicago increases its odometer reading by 150 miles*.

G540-2. The idea of "hill climbing with multiple restarts" (HC-MR) is that *N* times one chooses a random initial state and then performs standard hill climbing. The best solution over the *N* trials is returned as the result of the HC-MR algorithm.

Suppose you are given a state-space search problem specified by the nodes, legal actions, and a heuristic scoring function. Consider applying the HC-MR algorithm and a standard genetic algorithm (GA) to this search task.

(a) What do you feel is the fundamental difference between the HC-MR and GA approaches? Justify your answer.

(b) Discuss one relative strength and one relative weakness for each approach. (Do not reuse your answer to Part a.)

(c) Describe an experiment that you could run to empirically compare the performance of HC-MR and GA. Hypothesize under which conditions (i.e., properties of the given search task) each approach would perform best.

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. What is the shortest clause (fewest literals) that is consistent with the following positive and negative examples, given the background knowledge below? What solution would PROGOL return? What solution would FOIL return? (You should assume the basic FOIL algorithm, as described in the paper on the reading list and as discussed in CS731 and CS760, that at each step adds the literal with the greatest gain according to its scoring function.)

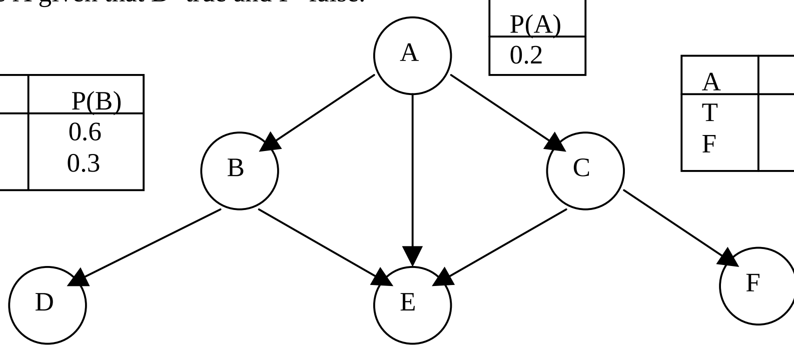
%% Positive examples.
p(a).
p(b).
p(c).

%% Negative examples.
:-p(d).
:-p(e).
:-p(f).

%% Background knowledge.
q(a,b).
q(b,c).
q(d,e).
q(e,f).
q(b,a).
q(c,b).
q(d,c).

B731-2. Given the following Bayes net, use variable elimination to compute the probability distribution over the variable A given that D=true and F=false.

P(A)
0.2



A	P(B)
T	0.6
F	0.3

A	P(C)
T	0.7
F	0.1

C	P(F)
T	0.5
F	0.9

B	P(D)
T	0.7
F	0.1

A B C	P(E)
F F F	0.2
F F T	0.1
F T F	0.3
F T T	0.1
T F F	0.5
T F T	0.8
T T F	0.6
T T T	0.9

B760 – Machine Learning Basic Questions

B760-1. The standard k-nearest neighbor (k-NN) approach is usually considered not to be a good choice of learning method for problems that may involve many irrelevant features.

(a) Why is standard k-nearest neighbor more sensitive to irrelevant features than say, decision tree induction algorithms?

(b) Sketch an approach to k-nearest neighbor learning that is better suited to problems with irrelevant features than standard k-NN.

(c) Discuss any disadvantages your approach has compared to standard k-NN.

B760-2. Consider the following training-set phrases involving single-letter "words"

<u>POSITIVE EXAMPLES</u>

A C D B
C C D D
A A B C
C D B D

<u>NEGATIVE EXAMPLES</u>

C A B D
A C D D
C D D C

And imagine that you wish to estimate to which category the following phrase belongs:

C C D A

(a) Apply the naïve Bayes algorithm to this task of assigning a category to the unlabeled example. Show all your work.

(b) Repeat (a), but this time consider the effect that the specific word at position i-1 has on the probability distribution for the word at position i (count words left-to-right).

(c) What would a decision-tree algorithm choose for its root node should it be given this training data? Be sure to justify your answer with explicit calculations.

B766 – Computer Vision Basic Questions

B766-1.

(a) Define, possibly using figures with your explanation, the terms (i) epipole, (ii) vanishing point, and (iii) disparity.

(b) Assume a parallel camera geometry with a baseline of 30cm between two identical cameras. Each camera has a focal length of 20mm and each pixel is 0.1mm x 0.1mm square. Compute the world coordinates of a point whose image coordinates (in pixels relative to the center of each image) are (20, 10) in the left image and (10, 10) in the right image. Assume that 1mm is 10 pixels.

B766-2.

(a) Describe the RANSAC algorithm in the specific case of fitting a line to a set of N points, 75% of which are expected to closely fit the line and the other 25% are outliers. Include descriptions of any parameters that are used and what appropriate values would be for this problem.

(b) Give some relative advantages and disadvantages of using RANSAC vs. the Hough transform for line detection.

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. Algorithms for inducing decision trees, neural networks, naïve Bayes classifiers, and first-order logic rules are all widely used in machine learning.

(a) Cite one advantage of each type of approach over the other three and justify your answers.

(b) For each of the four approaches, describe a particular application domain where that approach should be superior to the other three and explain why. (The application domain can be one you have read about or one that you formulate for the first time here.)

A760-2. There is a close relationship between minimum description length (MDL) approaches and Bayesian approaches to learning.

(a) Define and motivate the MDL principle.

(b) Describe a learning algorithm that uses the MDL principle. This can be an algorithm from the literature, or one of your own devising.

(c) Now state Bayes's rule and describe how various aspects of your learning algorithm relate to the terms in Bayes's rule.

A766 – Computer Vision Advanced Questions

A766-1.

(a) Sketch the visual motion field and the optical flow field for a rotating barber pole (i.e., a cylinder rotating about its major axis and on the cylinder is painted a strip that diagonally wraps around the cylinder several revolutions from top to bottom).

(b) How does the Factorization Method for recovering structure and motion deal with noisy data so as to produce a robust result? Be as specific as possible.

(c) Explain how the Factorization Method could be applied to a video sequence created by a fixed camera observing an object rotating 360 degrees on a turntable to produce a complete, closed surface representation of the object.

A766-2.

(a) Appearance-based object recognition using eigenspaces maps an input image into a k-dimensional eigenspace. Explain how this process can be thought of as an image compression technique. Give your answer in terms of a 512 x 512 image with 8 bits per pixel and k=10.

(b) What illumination assumption(s) is (are) required by the method, and why is this required?

(c) Explain how the method could be used to recognize 2D objects at any 2D position and 2D orientation in an image.