SPRING 1993
COMPUTER SCIENCES DEPARTMENT
UNIVERSITY OF WISCONSIN—MADISON
PH. D. QUALIFYING EXAMINATION

Artificial Intelligence
Depth Examination

Tuesday February 9, 1993
3:00 – 7:00 PM
113 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

- any two (2) questions chosen from questions 1–3 *and*
- any four (4) questions chosen from questions 4–11 *and*
- any two (2) questions chosen from questions 12–19.

Answer *eight* questions altogether.

Note: We are in a transitional year and the constraints will be tighter next fall for depth students. Students should check the Graduate Guidebook for information about the format of future exams.

POLICY ON MISPRINTS AND AMBIGUITIES:

The Exam Committee tires to proofread the exam as carefully as possible. Nevertheless, the exam sometimes contains misprints and ambiguities. If you are convinced 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.

All students must answer two (2) of the questions labeled B540/731.

B540/731 - 'Intro to AI' Questions

Question 1

Consider the following game.

> You and I take turns, with you going first. At each move, we must write down the number 5 or 7. The game lasts just four plies: you go, I go, you go, I go. At the end, we add up the four numbers played, and I pay you the sum of the numbers, modulo 10.

Draw the full game tree, and then explain how alpha-beta search works in searching the tree to determine your best opening move.

Question 2

For each of the following assertions, draw a *specific* search tree that supports it. Explain your answers. (Note: assume that best-first search solely uses *h*, the estimated distance to a goal state, to sort its OPEN list.)

a) Best-first search can explore more nodes than breadth-first search.

b) Hill climbing can produce a solution when depth-first search does not.

c) A* can produce a lower-cost solution path than does best-first search when both algorithms use the same admissible heuristic function *h*.

Question 3

Consider the following planning operators represented using situation calculus.

$\forall$ *Rm1,Rm2,S*	*connected(Rm1,Rm2)* $\wedge$ *in(robot,Rm1,S)*	$\rightarrow$	*in(robot,Rm2,do(move(robot,Rm1,Rm2),S))*
$\forall$ *Obj,Rm1,Rm2,Rm3,S*	*in(Obj,Rm1,S)* $\wedge$ *Obj* $\neq$ *robot*	$\rightarrow$	*in(Obj,Rm1,do(move(robot,Rm2,Rm3),S))*
$\forall$ *Obj,Rm,C,S*	*in(robot,Rm,S)* $\wedge$ *in(Obj,Rm,S)*	$\rightarrow$	*color(Obj,C,do(paint(robot,Obj,C),S))*
$\forall$ *Obj1,Obj2,C1,C2,S*	*color(Obj1,C1,S)* $\wedge$ *Obj1* $\neq$ *Obj2*	$\rightarrow$	*color(Obj1,C1,do(paint(robot,Obj2,C2),S))*

a) Add frame axioms to the above that say

(i) moving does not change color and

(ii) painting does not change location.

b) Assume the following describes the initial state of the world.

connected(a,b) $\wedge$ *connected(b,c)* $\wedge$ *in(robot,b,s0)* $\wedge$ *in(box,c,s0)* $\wedge$ *color(box,blue,s0)*

Derive a plan that achieves the goal below, reporting the binding for the variable *State*.

$\exists$ *State*   *color(box,red,State)*

(Note: *robot*, *box*, *blue*, *red*, *s0*, *a*, *b*, and *c* are constants, denoting two physical objects, two colors, the initial state, and three rooms, respectively.)

BREADTH students must answer any two (2) of the eight questions labeled B76x.

DEPTH students must answer any four (4) of the eight questions labeled B76x.

B760 - 'Machine Learning' Breadth Questions

Question 4

Assume you are given a two-category inductive learning problem, in which 100 training examples are specified using eight Boolean-valued features. Also assume that no two training examples have the same values for all eight features.

Imagine you decide to apply neural networks to this problem. Show that one layer of hidden units is sufficient to completely separate the training examples, regardless of the categorizations of the 100 examples. In the worst case, how many hidden units do you need for this specific task?

Question 5

a) Define the "overfitting" problem in machine learning and briefly discuss its significance.

b) How can it arise in

- decision trees
- neural networks
- explanation-based learning

c) For each of the three approaches listed in part (b), outline how one might use "tuning sets" (or "validation sets") to address the overfitting problem.

B761 - 'Deductive Prob. Solving & Logic Programming' Breadth Questions

Question 6

In Prolog, with the program

```
p :- r.
p :- q, not(r).
q :- not(r).
```

the query "`?- p`" will return an answer of "yes." What is the Clark completion of this program? Explain why `p` follows from it.

Question 7

Convert the following sentences to clauses, and then use the clauses to derive the empty clause, □, using resolution.

$$\forall x \, \forall y \, \forall z \quad (p(x,y) \Rightarrow p(x,z))$$
$$\forall x \, \forall y \, \forall z \quad (p(y,x) \Rightarrow p(z,x))$$
$$\exists x \, \exists y \, \exists u \, \exists v \quad (p(x,y) \wedge \neg \, p(u,v))$$

B762 - 'Knowledge Representation & Expert Systems' Breadth Questions

Question 8

Characterize the distinction between procedural representations of knowledge and declarative representations of knowledge. Give an example of a formalism for procedural representation and give an example of a formalism for declarative representation. In what sense and to what extent is it true that a Prolog program can be considered as a representation of both kinds at the same time? Compare the procedural-vs.-declarative distinction with the distinction between data and program.

Question 9

Expert systems are typically decomposable into a knowledge base and an inference engine. Illustrate this decomposition in terms of some one particular expert system (e.g., Mycin or Prospector), and indicate the primary features of the two components. What are the alleged advantages of such decomposability? Why not just embed the knowledge in the program code that constitutes the inference engine?

B766 - 'Computer Vision' Breadth Questions

Question 10

Consider the problem of recognizing three-dimensional blocks-world scenes, where the only types of allowable blocks are cubes and rectanguloids. That is, all faces are squares or rectangles, and all block vertices are formed from three faces meeting at right angles to one another. Hence this is a simplified world compared to the one used by Waltz. Derive what the junction dictionary for this simplified world should be. Show your work, including how you know the vertices in your dictionary are necessary and sufficient.

Question 11

Explain, being as specific as possible, why the Hough Transform algorithm is (a) good for detecting straight lines in "noisy" images, where some real points may be missing from a line and other extraneous points may be present elsewhere in the image, but (b) bad for detecting short line segments in large images containing lots of these line segments.

DEPTH students must answer any two (2) of the eight questions labeled D76x.

D760 - ‘Machine Learning’ Depth Questions

Question 12

Assume you have used backpropagation to independently train neural networks for classification tasks *A*, *B*, and *C*. You are next given task *D* to learn, for which tasks *A-C* are some of the subtasks. Describe how one could use the trained networks for tasks *A-C* when learning task *D*. Why would it make sense to try to do this? Discuss the strengths and weaknesses of the approach you propose.

(Feel free to answer this question by assigning concrete activities to tasks *A-D*. For example, task *A* might be “recognize doorways in a robot’s visual field.” Also, you need not use backpropagation to learn task *D*.)

Question 13

ID3 is a fast algorithm for producing decision trees; it is much faster than training neural networks via backpropagation. Given this, it seems appropriate to consider ways of modifying ID3 so that it uses more computer cycles before it chooses to output a decision tree as its answer.

Discuss how one could go about extending ID3 to make effective use of the plentiful computer cycles available these days. Feel free to consider 2-3 different approaches, and be sure to argue why your extensions have the potential to improve the generalization ability of ID3. Also outline any problems that might arise in the approach(es) you propose. Finally, describe experiments that would allow you to evaluate the effectiveness of your designs.

D761 - ‘Deductive Prob. Solving & Logic Programming’ Depth Questions

Question 14

If *f* is a binary function symbol, the only way that the equation

$$f(f(x,y),z) = y$$

can be valid is if the domain of discourse has only one element. Express this statement in resolution format, and then derive the empty clause using paramodulation plus binary resolution.

Question 15

With the Boyer-Moore prover, you can define the notion of “proper factor” by

```
(defn proper-factor (x y)
  (and (lessp 1 x)
       (lessp 1 y)
       (not (equal x y))
       (zerop (remainder y x)) ))
```

If you then try

```
(prove-lemma no-loops ()
  (not (and (proper-factor x y)
            (proper-factor y x) )) )
```

It fails. Why? What would you have to give it for hints and/or auxilliary lemmas to get this to work?

D762 - ‘Knowledge Representation & Expert Systems’ Depth Questions

Question 16

Describe and critique Minsky’s agent-based model of mind. Show how it attempts to account for several different cognitive phenomena like goal seeking, adaptability, and language understanding. How does Minsky avoid the challenge that any such model begs the question of explaining intelligence because its "society" of agents couldn’t possibly be intelligent unless the constituent agents were already intelligent? To what extent is it correct to identify the agents of the model with the objects of object- oriented programming (e.g., Minsky frames)?

Question 17

Symbolist theories of cognition are frequently challenged as failing to "ground" their symbols in what those symbols denote external to the organism. Explain the problem in terms of the SOAR "unified theory of cognition." How, if at all, does connectionism provide any help toward solving this problem? How, if at all, do AI models of vision (and/or of other kinds of perception) provide any help?

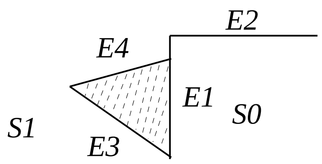
D766 - 'Computer Vision' Depth Questions

Question 18

Describe two alternative approaches to the problem of two-dimensional-object motion tracking, where one method is based on a spatiotemporal representation of the data, and the other is based on processing images frame by frame. Give an advantage and disadvantage of using each approach.

Question 19

Consider the computer vision "shape-from" problem of using shadow information to help recover surface orientations. Specifically, consider the following physical situation containing one planar surface, *S0*, occluding another planar surface, *S1*, with a shadow cast by *S0*'s two edges, *E1* and *E2*, onto *S1*, creating shadow edges *E3* and *E4*, respectively, due to an unknown point light source at infinity.



a) What are the unknown parameters that must be computed in this problem?

b) Derive *gradient space constraints* implied by the "virtual faces" corresponding to the two edges of *S0* that generate the two shadow edges on *S1*.