SPRING 1994
COMPUTER SCIENCES DEPARTMENT
UNIVERSITY OF WISCONSIN – MADISON
PH.D. QUALIFYING EXAMINATION

Artificial Intelligence
Depth Examination

Wednesday, February 9, 1994
3:00 – 7:00 PM
1257 Computer Sciences

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) of the questions B540-1, B540-2, and B540-3, *and*
- both questions in any *one* of the sections labeled B760, B761, B762, or B766, *and*
- any two (2) additional question in the sections B760, B761, B762, and B766, where these two questions need *not* come from the same section, *and*
- both question in the section labeled D76x that corresponds to the same area in which you answered both question in section B76x.

Hence, you are to answer a total of *eight* (8) 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 any two (2) of the three questions labeled B540-x.

B540 – Introduction to AI Questions

B540-1.

a. Define the *closed-world assumption* and discuss the role it plays in **Strips**.

b. Provide and explain a Strips-like representation of the operator `eat-lunch`. Discuss one aspect of this operator that was hard to capture, given the constraints of the Strips representation.

B540-2.

a. Represent the following in **first-order predicate logic**. Use the predicates $Purple(x)$, $Mushroom(x)$, and $Poisonous(x)$.

   i. There are no purple mushrooms.

   ii. All purple mushrooms are poisonous.

   iii. There is only one poisonous mushroom and it is purple.

b. For each of the following pairs of literals, determine if they unify. If so, give the most general substitution which unifies them; if not, give a brief explanation. x, y, and z are variable symbols, a and b are constants, f is a function symbol, and P is a predicate.

   i. $P(x, x)$ and $P(z, a)$

   ii. $P(a, f(y))$ and $P(x, x)$

   iii. $P(f(y), y, x)$ and $P(z, f(z), f(b))$

   iv. $P(z, z)$ and $P(f(x, y), f(y, x))$

B540-3. Explain how you would set up the traveling salesperson problem as a **best-first search**. You may assume that each instance of this problem will be a completely connected undirected graph with 30 vertices, where the cost of getting from one vertex to another is always some real number between 20 and 25. The problem is to output a cycle of minimal cost which visits each vertex. Describe what heuristic you will use, and whether or not it is admissible.

Answer BOTH of the question in any ONE of the sections labeled B76x. (Note: You must then answer BOTH of the questions in the corresponding section D76x.) Also answer any two additional questions in any of the other sections (these two questions need NOT occur in the same section).

B760 – Machine Learning Breadth Questions

B760-1. The number of hidden units in a **neural network** can greatly affect the network's ability to accurately categorize future examples.

- Explain why.
- Outline and discuss an experimental methodology for selecting a good number of hidden units for use with a given dataset.

B760-2. Decision-tree induction and nearest-neighbor classification are two successful ways to **learn from examples**.

- Discuss their relative strengths and weaknesses.
- Propose and justify a hybrid algorithm than combines aspects of both approaches.

B761 – Deductive Problem Solving & Logic Programming Breadth Questions

B761-1. Consider the **Prolog** program:

$q(a)$:– *not* $r(a)$.
$q(b)$.
$r(b)$.

The query ?– $q(a)$ will result in a *yes*, even though interpreting the program in the naive way, $q(a)$ doesn't follow; for example, the program doesn't comment on whether $r(a)$ is true or false. Explain what the *completion* of the program is and why $q(a)$ does follow from the completion.

B761-2. Convert the following **logical sentences** to clauses, and then derive the empty clause from the clauses using binary resolution plus paramodulation

$\forall xy(0 < y \;\Rightarrow\; x < x + y)$
$\forall xy(y < 0 \;\Rightarrow\; x + y < x)$
$\forall xy(x < y \vee y < x \vee x = y)$
$\forall x \neg(x < x)$
$\forall x(x = x)$
$\neg(\forall xy(x + y = x \;\Rightarrow\; y = 0))$

B762 – Knowledge Representation & Expert Systems Breadth Questions

B762-1. One popular architecture for expert systems emphasizes inference engine and knowledge base as the two primary components.

a. Clearly distinguish between the intended functions of the two components and give examples in terms of expert systems that have been well described in the literature (e.g., MYCIN).

b. Why is it hard to maintain the distinction when building efficient, real-world expert systems? (Hint: Note the difference between general inference rules and specialized, domain-specific inference rules.)

c. It has been argued that the distinction is useful only with respect to rule-based systems. Argue for or against its usefulness with respect to object-oriented systems.

B762-2. Suggest some of the ways that artificial neural nets might be used as components of expert systems. Specifically comment on how they might be used as an aid to the knowledge acquisition and formalization required for expert systems, and comment on whether the classical desideratum for expert systems, that they be able to "explain" their inferences, can be achieved in a system where neural nets are used for knowledge acquisition.

B766 – Computer Vision Breadth Questions

B766-1. The **Hough Transform** can be used to recognize geometric shapes in images. To detect a circle centered at coordinates (a, b) with radius r, we can parameterize it using the equation

$$(x-a)^2 + (y-b)^2 \quad = \quad r^2$$

Assume you are given a binary image produced by an edge detector such that a 1 indicates an edge point at the pixel's coordinates, and a 0 indicates no edge point was detected.

a. To detect circles with a radius of 5 pixels, describe what the Hough (i.e., parameter) space corresponds to. Include in your answer an explanation of what a point in this Hough space corresponds to in the original image.

b. If there are three co-circular points in the image (i.e., a circle of radius 5 passes through all three points), show how the voting process for these three points is related in the Hough space.

B766-2.

a. Describe the levels of representation computed in the **Marr Paradigm** for computer vision.

b. Give an example of a process that computes each of your levels in (a) from the previous level.

c. Explain how the "Principle of Least Commitment" is applied in the Marr Paradigm.

Answer BOTH of the questions in the section labeled D76x that corresponds to the same area in which you answered BOTH questions in section B76x.

D760 – Machine Learning Depth Questions

D760-1. Imagine you have used **reinforcement learning** to train a robot. Following training, you observe the behavior of the robot and notice several pieces of advice that you would like to give it. (For simplicity, assume the advice can be expressed as a collection of `IF-THEN` rules whose predicates directly refer to the robot's sensors and effectors.)

Present a design that would allow the robot to "take" this advice. That is, extend the basic reinforcement-learning approach so that it could make use of externally-provided, problem-specific knowledge. What happens if the advice is imperfect? Discuss any other strengths and weaknesses of your design.

D760-2. Discuss the claim that all (useful) **inductive biases** are a consequence of Occam's Razor.

D761 – Deductive Problem Solving & Logic Programming Depth Questions

D761-1. In the **Boyer-Moore prover**, `times` is built-in, so the following lemma:

```
(prove-lemma idem ()
     (implies
               (equal x (times x x))
               (or (zerop x) (equal x 1))))
```

is accepted as syntactically correct, but the prover fails to find a proof. Explain why, and then explain what auxiliary lemmas and/or hints you could give to the prover so that it would find a proof.

D761-2. Consider the following **logic program**, $\mathcal{P}$:

$a(0).$
$a(s(X)) :- \ a(X).$
$b(0).$
$b(s(X)) :- \ b(X).$
$p :- \ a(X), \neg b(X).$
$q :- \ \neg b(X), a(X).$
$same(Y, Y).$
$r :- \ same(X, s(X)).$

For each of the queries, $?- \ p$, $?- \ q$, $?- \ r$ answer the following:

a. Does it follow logically from the completion, $Comp(\mathcal{P})$?
b. Does its negation follow logically from $Comp(\mathcal{P})$?
c. How will a standard Prolog respond?
d. Is it true or false in the stratified negation model (or minimal model).

In all cases, explain why. A "standard Prolog" is one which does unsound negation and does not do an occur check.

D762 – Knowledge Representation & Expert Systems Depth Questions

D762-1

a. Characterize the distinction between knowledge represented procedurally and knowledge represented declaratively. Give examples of each. Briefly list the merits claimed for each.
b. Winograd appears to believe that the distinction is the same as the distinction between programs and data. Argue for or against his position.
c. Some so-called "formalists" argue that the distinction is the same as the distinction between inference rules and axioms/theorems in a formal logistic system. Argue for or against this claim.
d. Consider the representation of knowledge as objects (or frames). Explain why you consider this procedural representation, declarative representation, or both.

D762-2. Characterize the primary approaches used for representation of spatial knowledge in currently available "geographic information systems". What are the deficiencies of these approaches specifically with respect to their potential for generalization and for adding a temporal dimension, and how might they be improved in these respects by using more abstract data types (and perhaps other ideas associated with object-oriented knowledge representation)?

D766 – Computer Vision Depth Questions

D766-1.

a. Tomasi and Kanade's **Factorization method** solves for shape and motion without recovering depth from a sequence of images. Explain what this means, and what assumptions the method makes about the scene and/or imaging process.
b. Give a description of how this method works by decomposing the problem and the form of its solution method.

D766-2.

a. What is the **Active Vision** approach in computer vision? Give an example of a problem that has used this approach.
b. How might an active vision approach be used for object recognition, assuming that the information in the initial view is ambiguous? Briefly outline an algorithm, explaining in particular the parts that utilize the *active* capabilities you've chosen.