Doctoral Exam in Artificial Intelligence
May 1997

Answer all five problems. Each problem is worth 20 points.

Problem 1:

Let $\mathcal{L}$ be a first-order language over the domain of natural numbers containing the following primitives:

1, 6, 8 — Constants.
g(X, Y) — Predicate. $X > Y$.
f(X, Y) — Predicate. X is a factor of Y.
e(X, Y, Z) — Predicate. $X^Y = Z$.
p(X) — Predicate. X is prime.

A. Express the following statements in $\mathcal{L}$:

   i. For any P, if there exists $M > 1$ such that every factor of M is a power of P, then P is prime.
   ii. 6 is not prime.
   iii. $8 > 1$.
   iv. There is a factor of 8 that is not a power of 6.

B. Using resolution, prove (iv) from (i), (ii), and (iii). You should show the Skolemized form of all the clauses involved and each step of the resolution proof. You need not show the intermediate steps of Skolemization.

Problem 2:

I was at a workshop the other day, where an AI researcher showed a slide containing the following formula:

$$\text{on}(X, Y) \wedge \text{clear}(X) \wedge \text{clear}(Z) \wedge X \neq Z \wedge Y \neq Z \wedge \text{move}(X, Z) \Rightarrow$$
$$\text{on}(X, Z) \wedge \text{clear}(Y) \wedge \neg\text{clear}(Z) \wedge \neg\text{on}(X, Y)$$

This was supposed to mean, "If X is on Y and X is clear and Z is clear and $X \neq Z$ and $Y \neq Z$ and X is moved to Z, then X will be on Z and Y will be clear and Z will not be clear and X will not be on Y."

I tried to explain, politely, that this was totally wrong and a travesty of logical representations, but she was not convinced.

A. Explain, politely, what is wrong with the above formula.

B. Express the sentence correctly in a logical formalism. Briefly explain the meaning of your formalism.

Problem 3

Inductive inference is the problem of conjecturing a general rule from individual examples. One approach to inductive inference is Bayesian. Suppose we are interested in the rule R, "All α's are β's." We have chosen N α's at random, and we observe that they are all β's.

A. Let R be the event that the rule R is true. Let B_i be the event that the ith α selected satisfies β. Assume that

   - The prior probability of R is r.
   - The conditional probability of B_i given $\neg$R is p, and all the B_i are independent given $\neg$R.

   Of course, the probability of B_i given R is 1.

   What is the probability of R given $B_1 \ldots B_N$?

B. A key problem here is how to choose the prior probability r. What approach to this problem would the minimum description length theory of learning suggest for this?

Problem 4

Explain briefly (2 or 3 sentences each) each of the following terms from the theories of natural language understanding and computer vision.

A. Anaphoric ambiguity.

B. Chart parser.

C. Trigram model.

D. Stereopsis.

E. Segmentation.

Problem 5

A. A perfect-knowledge game with a random element, such as backgammon, can be represented using a tree with MIN nodes, MAX nodes, and nodes whose value is the expected value of their children. Explain.

B. Alpha-beta pruning in its usual form cannot be applied to the trees in (A). However, if there is a bound on the possible values of the evaluation function — for example, if it is known that the evaluation function is always between 0 and 1 — there are cases where an analogous pruning is possible. Give an example (you need not state the general rule.)