CMPSCI 683 Fall 2004 December 17, 2004

Final Exam

120 minutes, open book. For each question, explain your answer clearly and concisely.

Questions on Resource Bounded Reasoning (8 points, answer 1 out of 2)

RBR – 1: How can the current context of the environmental situation be used to tailor the "satisficing" response of a system in a resource-bounded situation. In our discussion of resource-bounded reasoning, an MDP was often used as as a control techniques for resource-bounded reasoning system. Why do you think that is?

RBR – 2: In the BB1 Guardian system, a number of ideas were integrated to achieve real-time responsiveness in time-critical situations: anytime knowledge sources/alternative KS for solving the same problem; a satisficing agenda management system; an adaptive, cognitively controlled I/O subsystem; a direct connection between the I/O subsystem and the affector subsystem; and a control architecture that permitted dynamic prioritization of specific KSs on a situation-dependent basis. Choose 3 of these and explain how they contribute to the system's real-time responsiveness.

Questions on Decision Theoretic Reasoning (15 points, answer 3 out of 4)

DTR –1: Explain the relationship between a decision tree and influence diagram. In what situations would an influence diagram be preferable?

DTR –2: The Markov assumption is that $P(s_t \mid s_{t-1}, s_{t-2}, \ldots, s_1, a) = P(s_t \mid s_{t-1}, a)$. Why does this significantly simplify the necessary information to represent a policy for an MDP? Why, for certain applications, is the size of the MDP state space very large to maintain the Markov assumption.

DTR –3: What type of uncertainty reasoning is possible in a Dempster-Shafer framework, yet more difficult in a classic probabilistic framework (e.g., a Bayesian Network)?

DTR –4: Why is defuzzification a necessary mechanism when using a fuzzy logic controller to make a decision?

Questions on Neural Networks/Rl (25 points, answer 5 out of the 7 questions)

NRL –1: Explain why RL will generally converge faster in environments which there are intermediate rewards, than they would in environments where all the rewards come from terminal states.

NRL –2: In explanation/analytic learning we saw how a domain theory could speed up classification learning. Would a domain theory help in RL, Neural Network or Decision-Tree learning? Choose 1 of these and give a more detailed explanation on how it would help.

NRL – 3: When Q learning is applied to non-deterministic actions the learning factor is based on how many times you have visited the state action pair. The more frequently a state action pair is visited the less importance is given each new training instance. Explain why this is important, and why this heuristic is not needed in the case where actions are deterministic.

NRL – 4: What are the differences and/or similarities between the credit assignment problem for neural networks and for that of reinforcement learning.

NRL –5: Neural Networks are supposed to have the following properties: noise resistance, error tolerance, graceful degradation, and learning with generalization. Why is this true?

NRL –6: In reinforcement learning there is a loop involving the use of an experience to refine the estimation of the values of state action pairs and then using these new values to recompute a new policy. Why is this loop key to an RL agent doing on-line learning?

NRL –7: Why did the developer of the champion backgammon playing system use a combination of Reinforcement Learning and Neural Networks in learning a policy for playing backgammon.

Questions on Other Aspects of Learning (15 points, answer 3 out of 5)

LEA – 1: If in instance-based learning you use a distance-based metric for weighing the contribution of a training instance to the classification of a new instance, why does it then make sense to use the entire set of training instances rather than just the K nearest ones?

LEA –2: Why is conceptualizing search as an integral part of the process of learning a reasonable viewpoint? What type of search is used in neural networks? How would you compare that type of search with the one used in the version space algorithm?

LEA –3: What causes overfitting in a decision classification tree? Give some examples of techniques to correct overfitting problems for this type of learning.

LEA –4: Is there a relationship between the concepts of overfitting and learning-bias to probably approximately correct (PAC) learning? If you think there is, explain the relationship. Why is PAC not always a good estimator of the effectiveness of a specific learning strategy?

LEA –5: Explain the unique challenges of relational learning. What aspects of the input data are different from those in many traditional learning tasks? What opportunities can relational learning algorithms exploit that many traditional learning algorithms cannot?

Decision Theoretic Reasoning (21 points)

This is an example discussed in class that has been slightly modified to make it more interesting!

Suppose an oil company is hoping to buy one of *n* blocks of ocean drilling rights.

- Exactly 2 blocks contain oil each worth *C* dollars.
- The price of each block is 2*C/n* dollars.
- A seismologist offers the company a survey indicating whether block #3 contains oil. He is right 50% of the time.

How much should the company be willing to pay for the information?

MDPs (16 points)

The following is an example we discussed in class. On the left is the policy of where to move at each box in order to maximize reward. On the right is the expected reward that you will achieve if you are in a specific box.

Actions succeed with probability 0.8 and *move at right angles with probability 0.1* (remain in the same position when there is a wall). Actions incur a small cost (0.04).

Why move to box .655 instead of box .660 from Box .611? Show the calculation. (10 points)

What happens when cost increases, does the policy change? (6 points)