Answer THREE questions, at least ONE from EACH of Sections A and B.

SECTION A

1. a. Give your answers to (i)-(iv) briefly.

   i. What is a non-monotonic reasoning system?

   ii. Why is monotonicity desirable in the reasoning behaviour of an intelligent agent?

   iii. Why should an intelligent agent want to suspend monotonicity?

   iv. What is the qualification problem, and why is it important?

   [10 marks]

   b. Suppose you are working as an artificial intelligence consultant. Your client is a manufacturer of a new type of electrical generator that is used in remote locations. The client needs an intelligent diagnostic system for use by mechanics in the field for fixing faults in the electrical generator. Furthermore, suppose your client is considering either a Truth Maintenance System or a Bayesian Network System. Give a short explanantion of each approach and of how it could be used in this application. What are the pros and cons of each of these options, and what would your recommendation be?

   [13 marks]

   c. Consider the following default theories. Give the extensions, with explanations, if there any, otherwise explain why no extension exists.

   i.

$$(\{\frac{\alpha \wedge \beta : \gamma}{\delta}, \frac{\beta \wedge \delta : \neg\gamma}{\alpha}\}, \{\beta\})$$

   ii.

$$(\{\frac{\alpha : \beta}{\gamma}, \frac{\alpha : \neg\gamma}{\neg\beta}\}, \{\neg\alpha \rightarrow \alpha\})$$

   [10 marks]

[Total 33 marks]

2. a. Define some features that could be used to informally delineate the notion of an autonomous intelligent agent.

[5 marks]

b. Consider the following patterns of relevance between random variables. For each of the patterns, give a brief explanation and an example.

i. Marginal and conditional independence.

ii. Marginal and conditional dependence.

iii. Marginal independence and conditional dependence.

iv. Marginal dependence and conditional independence.

[10 marks]

c. What is the main computational problem with Bayesian networks and how can it be addressed?

[6 marks]

d. i. Give a definition for discrete linear time models for temporal logic.

ii. Give a definition for the semantics for a tense logic. State the tense logic you have chosen.

[12 marks]

[Total 33 marks]

3. a. For Dempster-Shafer theory, define the following concepts:

   i. A frame of discernment

   ii. A basic probability assignment

   iii. A belief function

   iv. A plausibility function

[10 marks]

b. What is the Dempster rule of combination? Give a definition and discuss some of the pros and cons of it.

[10 marks]

c. The following two basic probabilities assignments are defined for the frame of discernment = $\{\alpha, \beta, \gamma\}$:

$$
\begin{array}{ll}
m_1(\{\alpha\}) = 0.5 & m_2(\{\alpha\}) = 0.0 \\
m_1(\{\beta\}) = 0.0 & m_2(\{\beta\}) = 0.2 \\
m_1(\{\gamma\}) = 0.0 & m_2(\{\gamma\}) = 0.3 \\
m_1(\{\alpha, \beta\}) = 0.3 & m_2(\{\alpha, \beta\}) = 0.0 \\
m_1(\{\alpha, \gamma\}) = 0.0 & m_2(\{\alpha, \gamma\}) = 0.0 \\
m_1(\{\beta, \gamma\}) = 0.0 & m_2(\{\beta, \gamma\}) = 0.5 \\
m_1(\{\alpha, \beta, \gamma\}) = 0.2 & m_2(\{\alpha, \beta, \gamma\}) = 0.0
\end{array}
$$

Give the combined basic probability assignment, the resulting belief function, and the resulting plausibility function. Explain your calculations.

[13 marks]

[Total 33 marks]

SECTION B

4. a. Give a detailed definition of what it means for a hypothesis space H to be 'potentially learnable'.

[12 marks]

b. Prove that if H is finite then it is potentially learnable.

[18 marks]

c. What does the previous result allow you to say about a consistent learning algorithm for (H, H) when H is finite.

[3 marks]

[Total 33 marks]

5. a. Outline a procedure that can be used to construct a decision tree. Include in your answer a definition of 'gain', and an outline of the way in which it can be used.

[12 marks]

b. Explain what is meant by the term 'overfitting', and give an example of what can cause it. Outline the effect it generally has. Give an example of a technique that can be used to try to prevent overfitting when producing a decision tree.

[6 marks]

c. Let A be an algorithm that allows us to produce a decision tree from any sequence $\mathbf{s}$ of training examples. Explain how the error probability of the decision tree produced using all the examples in $\mathbf{s}$ can be estimated, first using the resubstitution estimate, and second using the cross-validation estimate. Compare the general effectiveness of these two techniques.

[15 marks]

[Total 33 marks]

6. Write an essay, including as much technical detail as you can, on the subject of *Probably Approximately Correct (PAC) Learning*.

[Total 33 marks]

END OF PAPER