COMP34120

Two hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

AI and Games

Date: Friday 18th May 2012

Time: 09:45 - 11:45

Please answer any THREE questions from the FOUR questions provided

Use a SEPARATE answerbook for each Question

Marks will be awarded for reasoning and method as well as being correct.

This is a CLOSED book examination

The use of electronic calculators is NOT permitted

[PTO]

1. a) Alice and Bob play the following game with a deck of cards consisting of one Queen and one King. Each pays one pound into the pot.

   Alice draws a card and looks at it without showing it to Bob. She can now either bet or fold. If she folds, Bob gets the pot and the game is over. If she bets, she has to add another pound to the pot.

   Bob can now either fold, in which case Alice gets the pot and the game is over, or he can accept her bet by also putting another pound into the pot, or he can double the existing bet by putting two more pounds in the pot. If he doubles then Alice has to put another pound into the pot.

   Now Alice turns over her card. If she has a Queen then Bob gets the pot, otherwise Alice gets it.

   i) Draw a game tree for this game. (4 marks)

   ii) Describe all strategies for both players. (4 marks)

   b) Consider the following game for two players:

$$\begin{vmatrix} (4,4) & (4,4) \\ (5,5) & (4,4) \end{vmatrix}$$

   Determine the pure equilibrium points of the game. What can you say about their desirability for the two players? How do you think the game will be played in practice? (3 marks)

   c) Describe as many methods as you can think of for calculating the equilibrium points for a 2-person zero-sum game. (6 marks)

   d) Give a 2-person zero-sum game in extensive form of size (2×2) with at least three different entries which has precisely two pure strategy equilibrium points. Is it possible to give such a game with four different entries? Justify your answer. (3 marks)

2. Consider the game Kalah, with seven wells per player and seven stones in each well.

   a) How would you classify this game? What can you say about its possible solutions? (4 marks)

   b) Give an overview of the program your group wrote to play Kalah. What were the main parts of the program and how did they fit together? (4 marks)

   c) What do you consider to be the greatest strength of your group's program? How do you think that affected your program's performance? Justify your answer. (4 marks)

   d) What do you consider the greatest weakness of your group's program? How do you think that affected your program's performance? Justify your answer. (4 marks)

   e) What would have been the one idea that could have considerably strengthened your program? Justify your response. (4 marks)

3. a) Describe what are two person Stackelberg games with perfect information and imperfect information respectively. Further explain what is the main difference when solving Stackelberg games with perfect information and imperfect information. (4 marks)

   b) Find a Stackelberg strategy for the following Stackelberg game:

   - There are two players in which Player L is the leader and Player F is the follower;
   - The strategy spaces for the leader and the follower are $U_L = [0, +\infty)$ and $U_F = [0, +\infty)$;
   - The payoff functions for the leader and the follower

   $$J_L(u_L, u_F) = (u_L - 1)(3 - 2u_L + u_F)$$
   $$J_F(u_L, u_F) = -{u_L}^2 + 2u_L u_F - {u_F}^2 - 4u_L + 4u_F + 5$$

   in which $u_L \in U_L$ is the leader's strategy and $u_F \in U_F$ is the follower's strategy.

   Note. 8 marks are for the step by step process to obtain the solution, 1 marks for the correct answer. (9 marks)

   c) In two person Stackelberg games with imperfect information, if the environment is time-varying, the follower's reaction function needs to be learned and kept updating. Within this context, answer the following questions:

   i) What is the moving window approach for learning and updating the follower's reaction function? (2 marks)

   ii) What is the weighting least square method with forgetting factor for learning and updating the follower's reaction function? (2 marks)

   iii) Give the formula of recursive least square method and explain the meaning of each item in the formula (Note. When giving the formula, you do not require to give the detailed formulas for updating the adjusting factor and initial conditions). (3 marks)

4. a) What is reinforcement learning? How does reinforcement learning differ from supervised learning and unsupervised learning as studied in COMP24111?
(4 marks)

b) Why is reinforcement learning relevant to learning in games? Give another (non-game) situation in which reinforcement learning can be used.
(2 marks)

c) In reinforcement learning, often one wants to find a solution which optimises some criterion.

i) One possible criterion to optimise is the "regret". Give a definition of regret, both in words and mathematically. (3 marks)

ii) Give another criterion which one might want to optimise. (1 mark)

d) Recall the prisoner's dilemma, with a payoff as follows:

		Player 2 C	Player 2 D
Player 1	C	$-2,-2$	$-10,0$
	D	$0,-10$	$-8,-8$

You are going to make a learning bot to play a repeated version of this game, repeated for 200 games, against a community of other bots. You don't know how the other bots work (and you cannot reverse engineer them) but you can download them and play your bot against them.

i) If you could compute the Nash equilibrium, should you play it? Why or why not? (2 marks)

ii) Can min-max search with alpha-beta pruning be used to solve this? Why or why not? (2 marks)

iii) Describe, based on methods learnt in this course, how you would make a learning bot to learn to play well against this community of bots. Be sure to describe how you would represent the bot's strategies, how new policies would be explored, and how strategies would be updated during learning based on the outcomes of the game. In order to make a strong player, the strategies may depend on the history of the opponent in the previous play or plays (such as tit-for-tat, which depends on the play of the opponent in the previous play).
(6 marks)

END OF EXAMINATION