COMP 34120

Two hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

AI and Games

Date: Thursday 16th May 2019

Time: 09:45 - 11:45

Please answer all FOUR Questions
Each question is worth 15 marks

Use a SEPARATE answerbook for each QUESTION

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

[PTO]

Section A

Semester 1

1. Normal Form Games

a) What follows is the pay-off for a two-player, zero-sum game with perfect information and no chance. Player 1 has three actions: A – C. Player 2 has four actions: a – d. The table shows the pay-off to Player 1; the pay-off to Player 2 is minus this. For example, if Player 1 plays B and Player 2 plays a, Player 1 gets 4 and Player 2 gets -4. The two players choose their moves simultaneously.

		Player 2			
		a	b	c	d
Player 1	A	5	6	1	2
	B	4	2	4	2
	C	2	3	3	1

i) Find a (Nash) equilibrium using the minimax approach. Justify your answer. (3 marks)

ii) The concept of *dominance* can be used to reduce the size of the pay-off table in certain instances. Find a (Nash) equilibrium by repeatedly applying dominance to shrink the pay-off table as much as possible. Make clear the path of row and column reduction to get the equilibrium. (3 marks)

b) This question involves the following game.

The Malicious Diner "Game" There is a restaurant which offers only two choices: 2-course set menu for £20 and 3-course set menu for £30, drinks and service (tip) included.A group of people go there regularly, and agree to split the bill evenly, regardless of what was ordered. Ordering is done on a tablet, so what a person ordered is only known to that person and the waiter until the food arrives. The malicious dinner orders the more expensive meal hoping to get it a a reduced price due to other diners who have ordered the cheaper one.

This question concerns the two diner version. Each diner chooses one of the two meals, and the choice is unknown to the other diner. The payoff to each diner is the difference between the value of the meal that diner received and the amount it paid. For example, if Alice orders the £30 meal and Bob orders the £20, they each pay 25 (since they are splitting the bill evenly) and Alice gets a reward of £30 – £25.

i) Convert this into a game in normal form, i.e. write the game table. What type of game is this? (3 marks)

ii) Find the pure strategy equilibrium. If there is more than one, find all of them. For every equilibrium found, justify that it is a Nash equilibrium. (3 marks)

iii) Is there a fully mixed strategy equilibrium? If so, find it. If not give a convincing argument that one does not exist. (3 marks)

2. Extensive Form Games

a) Consider Kalah $(2, 1)$, that is Kalah with two pits per side and one stone in each pit to start the game. Figure 1 shows the starting position of the game. Consider the case with the pie rule. Solve Kalah $(2, 1)$ with the pie rule. What level of solution is your solution? (5 marks)

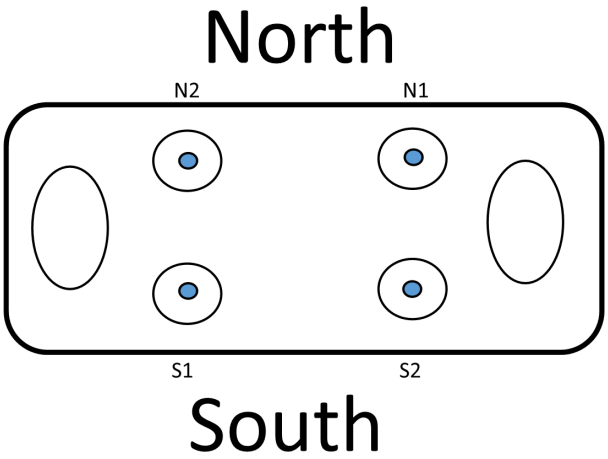

Figure 1: The starting position for $(2, 1)$ Kalah. South is player 1 and goes first.

Too hard? If you are stuck and cannot solve the game, for 2 of the 5 marks write down all of the possible answers; what are the possible ways a game like this can be solved?

The rules of Kalah: In case you need a reminder of the rules of Kalah

i. South is player 1 and goes first. On the first move, South gets only one turn even if its last deposited seed is in its scoring well.

ii. At North's first turn, it can either play as normal or SWAP. If North SWAPS, North becames South, South becomes North, and it is North's turn.

iii. The player whose turn it is chooses one of the 2 holes they own which is not empty. The seeds from that hole are removed and deposited one by one into the subsequent holes moving counter clockwise, skipping only the opponent's scoring well.

iv. If the last seed is deposited into the current player's scoring well, that player must take another turn.

v. If the last seed is deposited into an empty hole owned by the current player and the hole opposite is not empty, the player captures all of

the seeds in the opposite hole plus the capturing seed, which go into the player's scoring well.

vi. The player's turn ends with a capture or when the last seed is deposited into a hole other than the player's scoring well.

vii. If a player is on turn and has no legal move (e.g. all its holes are empty), the game ends. All seeds not in one of the scoring wells are deposited in the scoring well of the opposing player. The player with the most seeds in their scoring well wins.

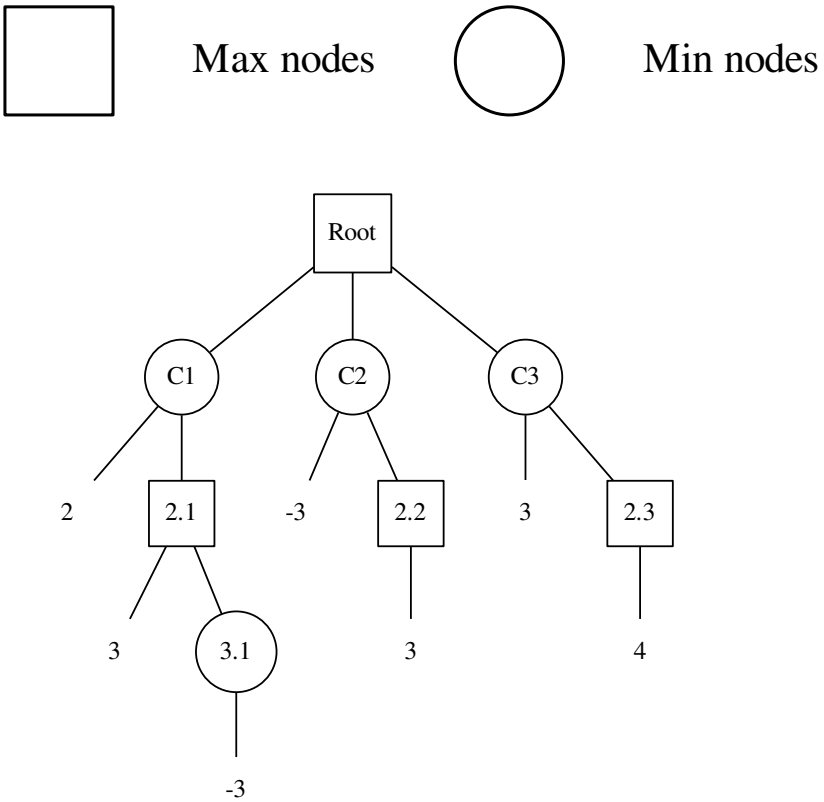

Figure 2: A game tree for Question 2b. Square nodes are Player 1 (MAX) decision nodes. Circle nodes are Player 2 (MIN) decision nodes. Numbers without shapes are payoffs. The numbers within the shapes are labels.

b) Figure 2 shows a game tree in extensive form. Apply alpha-beta pruning to this tree to answer the following questions:

   i) After the algorithm is completed, what are the alpha, and beta values of the root node and its three children? (4 marks)

   ii) What nodes are *not* evaluated? (2 marks)

c) These questions are about heuristics.

   i) Describe two heuristics for Kalah. Be sure describe the heuristic in a way which makes clear how the heuristic assigns a numerical value to a board position. (2 marks)

   ii) The stronger your heuristic, the stronger your player will be. Describe the most effective technique you know of to build a strong heuristic. Describe also how you would compare heuristics to determine which are the most effective ones. (2 marks)

Section B

Semester 2

3. a) Consider the following Stackelberg game:

- There are two players in which L is the leader and F is the follower, and the strategy spaces for the leader and the follower are U_L and U_F respectively;
- The payoff functions for the leader and the follower are

$$J_L(u_L, u_F) = -(3u_L - u_F)^2 - 6u_L + 2u_F - 1$$
$$J_F(u_L, u_F) = -4u_L^2 + 4u_L u_F - u_F^2 - 4u_L + 2u_F - 1$$

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

Assume that the follower's strategy space is a continuous one given as $U_F = [0, +\infty)$. Now for each of the following cases, find a Stackelberg strategy:

i) The leader's strategy space $U_L = [-1, +\infty)$. (6 marks)

ii) The leader's strategy space $U_L = [1, +\infty)$. (2 marks)

Note. For each case above, 1 mark is for the correct answer and the other marks are for the step by step calculation and reasoning.

b) Consider the following Stackelberg game problem:

- There are two players, Player 1 and Player 2, whose strategy spaces are $U_1 = (-\infty, +\infty)$ and $U_2 = (-\infty, +\infty)$ respectively;
- The payoff functions for Player 1 and Player 2 (no matter whether they are leader or follower) are

$$J_1(u_1, u_2) = -(u_1 - u_2)^2 + 6u_1 - 9$$
$$J_2(u_L, u_F) = -(u_2 - 2u_1)^2$$

in which $u_1 \in U_1$ is Player 1's strategy and $u_2 \in U_F$ is Player 2's strategy.

Now answer the following questions:

i) Assume that Player 1 has the choice to be either the leader or the follower. Then is there a choice which is better off for Player 1? If so, which one? In any case, explain why? (5 marks)

ii) Assume that Player 2 has the choice to be either the leader or the follower. Then is there a choice which is better off for Player 2? If so, which one? In any case, explain why? (2 marks)

Note. For each question above, 1 mark is for the correct answer and the other marks are for the detailed analysis and reasoning.

4. a) In a two-person Stackelberg game with imperfect information, the follower's reaction function $y = R(x)$ can be learned from historical data $\{y(t), x(t)\}$ $(t = 1, 2, ..., T)$. Assume that the best estimated reaction function $y = \hat{R}(x, \theta)$ is learned from minimising the sum of square errors. Now answer the following questions:

   i) Please give two error functions which can be used to measure the quality of the learned reaction function $y = \hat{R}(x, \theta)$ and briefly discuss their differences. (4 marks)

   ii) Now assume that the follower's reaction function $y = R(x)$ is time-varying, and the best estimated reaction function $y = \hat{R}(x, \theta_T)$ is learned by using the moving window approach with T_w as the width of the moving data window. How to modify the above two error functions to measure the quality of the learned reaction function $y = \hat{R}(x, \theta_T)$. (2 marks)

   b) Consider the following mechanism design problem of a multi-item auction game: There are two items: Items 1 and 2 which include 100 and 50 identical products respectively. There are three bidders, Bidders 1, 2, and 3, whose private valuations per product are 20, 10 and 5 respectively. Assume that the objective of bidder $i (i = 1, 2, 3)$ is to maximise his utility $u_i = v_i * x_i - p_i$ $(i = 1, 2, 3)$, where v_i is bidder i's valuation, x_i is the number of the product within an item if bidder i wins the item and 0 otherwise, and p_i is the price paid to an item if bidder i wins the item. The sealed-bid auction has the following three steps: 1) Collect bids $b = (b_1, b_2, b_3)$; 2) Allocate the items to the bidders based on a designed allocation rule in which each bidder can get at most one item); 3) Charge each bidder who wins an item based on a designed payment rule. Further assume that the auctioneer's objective is to maximise the social surplus function $\sum_{i=1}^{3} v_i * x_i$. For the above auction game, answer the following questions:

   i) Please give the definitions that 1) a designed mechanism is Dominant Strategy Incentive Compatible (DSIC); 2) an allocation rule is monotone; 3) an allocation rule is implementable. (3 marks)

   ii) For the above auction, design an allocation rule and a payment rule to make the resulting mechanism be DSIC. (3 marks)

   iii) Under your designed allocation and payment rules above, how many products each bidder will get and how much each bidder will pay? Further what is the max value of the social surplus function achieved? (3 marks)

END OF EXAMINATION