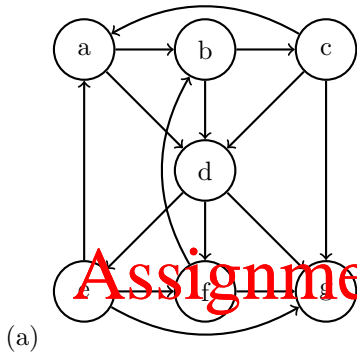


COMP4418, 2017 – Assignment 3

Due: Wednesday, 22 November, 23:59:59

Worth: 15%

1. [20 Marks] (Social Choice and Game Theory)



In the tournament
downward are

- the uncovered set;
- the top cycle;
- the set of Copeland winners;
- the set of Banks winners; and
- the set of Condorcet winners.

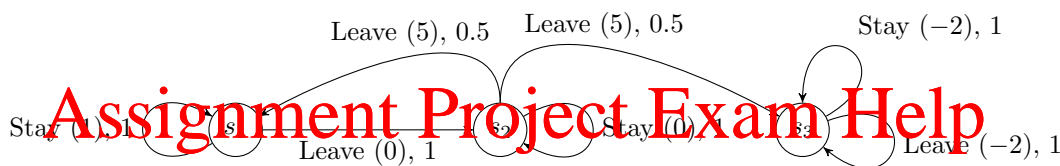
(b) Compute all the Nash equilibria of the following two player game.

| | D | E |
|---|------|------|
| A | 2, 4 | 8, 5 |
| B | 6, 6 | 4, 4 |

2. [30 Marks] (Decision Making)

- (a) For each of the following games, choose the best model among (A) Markov chain (Markov process); (B) Markov decision process (MDP); (C) Hidden Markov model (HMM); (D) Partially-observable Markov decision process (POMDP); and (E) None/Other.
- Blackjack
 - Candy Crush
 - Chess
 - Minesweeper
 - Snakes and Ladders
 - Texas Hold 'em Poker

For the next questions, consider the Markov Decision Process depicted below. Edges are labelled “name of the action (reward associated), probability of the transition”.



- (b) Using your intuition, give an optimal policy for situations where the value is high (for instance, $\delta = 0.001$). Explain your reasoning.
- (c) Using your intuition, give an optimal policy for situations where the value is low (for instance, $\delta = 0.001$). Explain your reasoning.
- (d) Represent the values computed during the first three iterations using the following format where L represents the action *Leave* and S represents the action *Stay*. Use a discounting factor of 0.6.

| | $V_0(s)$ | $V_0(s, S)$ | $V_0(s, L)$ | $V_1(s)$ | $V_1(s, S)$ | $V_1(s, L)$ | $V_2(s)$ | $V_2(s, S)$ | $V_2(s, L)$ | $V_3(s)$ |
|-------|----------|-------------|-------------|----------|-------------|-------------|----------|-------------|-------------|----------|
| s_1 | 0 | 1 | ... | | | | | | | |
| s_2 | 0 | ... | | | | | | | | |
| s_3 | 0 | | | | | | | | | |

- (e) Let π be the following policy: $\pi(s_1) = L$, $\pi(s_2) = L$, $\pi(s_3) = S$. If π is assumed to hold, the MDP turns into a Markov Chain. Represent this Markov Chain / Markov Process.
- (f) Assuming the agent uses π , express the value associated to each state as a function of the discount factor δ . Provide the formal derivation of the result as part of your answer. Elaborate on whether the computations of this question support the intuition of questions 2b and 2c.

Submission

- Put your written solutions in a single PDF file `assn3.pdf`
- Submit using the command: `give cs4418 assn3 assn3.pdf`

Late Submissions

Due to the assignment due date being extended to the 22nd November there will be no late submissions allowed.

Academic Honesty and Plagiarism

All work submitted for assessment **must be your own work**. Assignments **must be completed individually**. We regard copying of assignments, in whole or part, as a very serious offence. Be warned that:

- the submission of work derived from another person, or jointly written with someone else will, at the very least, result in automatic failure for COMP4418 with a mark of zero;
- allowing another student to copy from you will, at the very least, result in a mark of zero for your own assignment; and
- severe or second offence possibly other academic discipline

Collaborative work in the form of “think tanking” is encouraged, but students are not allowed to derive solutions together as a group during such discussions. Students are also warned not to share fragments of the assignments to each other in any form (e.g. as email or listings). In a similar way, selling/purchasing of solutions that is available on the web is also not permitted. Students are advised to protect their work. Do not leave your terminal/computer unattended. Do not leave the printer for others to take. Read the study guide carefully for the rules regarding plagiarism.