CPSC 532L: Models of Strategic Behaviour **Spring 2023**

Assignment #5

Due: April 23rd, 2023, 1:00pm

Instructor: Kevin Leyton-Brown *email: narunram@cs.ubc.ca*
TA: Narun Raman

1 Finitely repeated games

Problem 1.1. **[6 points]** Consider the game "chicken," repeated four times:

		P_2: G	P_2: S
P_1	G	$(-10, -10)$	$(3, -2)$
	S	$(-2, 3)$	$(-1, -1)$

(a) **[3pts]** Find a subgame perfect equilibrium of this game where both players get utility of 2.

(b) **[3pts]** Find a subgame perfect equilibrium of this game where crashing (G, G) happens with positive probability in every round.

2 Folk theorem, repeated games, and correlated equilibrium

Problem 2.1. **[20 points]** Consider the following extensive form game:

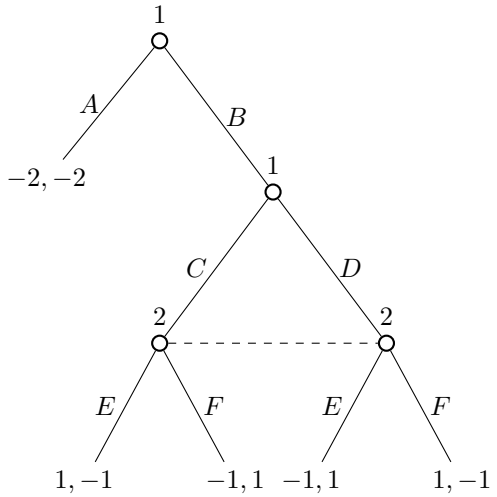


Figure 1: Extensive-form game

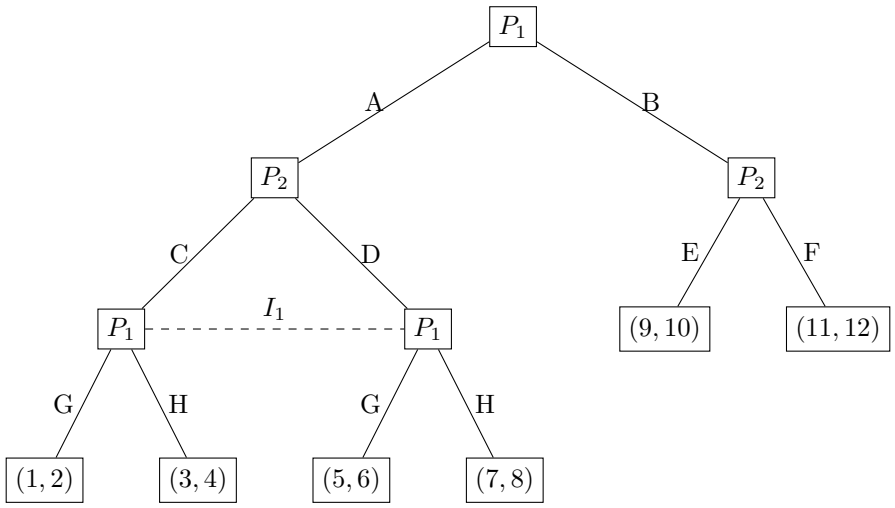
(a) **[4pts]** For the unrepeated game, what are the maxmin strategies and maxmin values for each player?

(b) **[6pts]** Prove or disprove the following statement: "For the infinitely repeated game, there exists a Nash equilibrium where player two's (limit-of-the-means) average reward is -2."

(c) **[10pts]** Construct a correlated equilibrium of the infinitely repeated game where player one's average reward is 1/4. Prove that your solution is a correlated equilibrium. (Note that any Nash equilibrium of an infinitely repeated game is also a correlated equilibrium of that game.)

3 Sequence Form

Problem 3.1. [10 points] Consider the following game:



(a) **[1pts]** What is the induced normal form of this game?

(b) **[1pts]** What are the sequences for P_1 and P_2?

(c) **[2pts]** What is the payoff matrix for the sequence form of this extensive form game?

(d) **[2pts]** What advantage does the sequence form matrix have over the INF matrix of (a) and how might this help with computations using it?

(e) **[1pts]** Find the equivalent realization plan r_1 of the behavioral strategy of P_1 taking either of their actions at each of their information sets with $\frac{1}{2}$ probability.

(f) **[3pts]** Write the objective of the primal linear program that computes P_1's best response and the state in words what linear constraints must be obeyed in the primal LP assuming we are given P_2's realization plan. What feature of the primal objective would make us want to use the dual LP instead, if we additionally had to optimize with respect to P_2's realization plan to find an equilibrium?

4 Diplomacy and CICERO

Problem 4.1. [10 points]

(a) **[2pts]** Is Diplomacy a perfect information or an imperfect information game? Explain your answer.

(b) **[3pts]** The Dialogue Module of CICERO is conditioned on *intents*. Remember from the presentation that a message is defined to have *intent* $\mathbf{z}$ if $\mathbf{z}$ is the most likely set of actions the sender and recipient will take. The Dialogue Module was finetuned through supervised learning, where it received inputs consisting of the board state, dialogue history and intents, and it was taught to predict a message. The training data for the board state, dialogue history and the output message were all taken from a large dataset of Diplomacy games, while the intents were artificially generated. Why did the authors of the

paper have to generate the intents artificially, and could they not have extracted the intents from the actions that were taken in each round? *Hint: think about how the generated intents and conversations may differ from the actions that are taken in the game.*

(c) **[2pts]** What is a limitation of conditioning the Dialogue Module on *intents*?

(d) **[3pts]** The Strategic Reasoning Module of CICERO models other players' strategy using the following utility function:

$$U_i(\pi_i, \pi_{-i}) = u_i(\pi_i, \pi_{-i}) - \lambda D_{KL}(\pi_i || \tau_i),$$

where π_i is a player's strategy, π_{-i} is the strategy of all other players, τ_i is the anchor policy, and $u_i(\pi_i, \pi_{-i})$ is the expected value based on self-play RL. What is the impact of λ in this formula (i.e. what does it mean for λ to have a large and small value) and how does the value of λ influence the extend to which the dialogue history determines the policy? *Hint: think about which of these terms are influenced by the dialogue history.*

Academic Honesty Form

For this assignment, it is acceptable to collaborate with other students provided that you write up your solutions independently. The only reference materials that you can use are the course notes and textbook, and the reference textbooks listed on the course web page. In particular, getting help from students or course materials from previous years is not acceptable.

List any people you collaborated with:

1.

2.

3.

List any non-course materials you refered to:

1.

2.

3.

Fill in this page and include it with your assignment submission.