CS 798: Multiagent Systems

Repeated and Stochastic Games

Kate Larson

Computer Science
University of Waterloo

Outline

Introduction

- So far we have assumed that a normal-form game is played once, but there are many situations where this is not true
  - The normal-form game is called a *stage game*
- When repeating a game there are new questions that need to be addressed
  - What can agents' observe?
  - What can agents' remember?
  - How do you define agents' utility?
- The answers to these questions can depend on whether a game is repeated a finite or infinite number of times.

Introduction

- So far we have assumed that a normal-form game is played once, but there are many situations where this is not true
  - The normal-form game is called a *stage game*
- When repeating a game there are new questions that need to be addressed
  - What can agents' observe?
  - What can agents' remember?
  - How do you define agents' utility?
- The answers to these questions can depend on whether a game is repeated a finite or infinite number of times.

Introduction

- So far we have assumed that a normal-form game is played once, but there are many situations where this is not true
  - The normal-form game is called a *stage game*
- When repeating a game there are new questions that need to be addressed
  - What can agents' observe?
  - What can agents' remember?
  - How do you define agents' utility?
- The answers to these questions can depend on whether a game is repeated a finite or infinite number of times.

Finitely Repeated Games

Play the Prisoners' Dilemma several times with a partner?
What happens?

	C	D
C	-1, -1	-4, 0
D	0, -4	-3, -3

$\Rightarrow$

	C	D
C	-1, -1	-4, 0
D	0, -4	-3, -3

Finitely Repeated Games

- Finitely repeated games can be modeled as extensive form games
- They have a much richer strategy space
  - A *stationary* strategy has an agent playing the same strategy in each stage game (memoryless)
  - In general, however, strategies can depend on previous actions
- If the stage game has a dominant strategy then you can solve the finitely repeated version by applying backward induction

Finitely Repeated Games

- Finitely repeated games can be modeled as extensive form games
- They have a much richer strategy space
    - A *stationary* strategy has an agent playing the same strategy in each stage game (memoryless)
    - In general, however, strategies can depend on previous actions
- If the stage game has a dominant strategy then you can solve the finitely repeated version by applying backward induction

Infinitely Repeated Games

- Consider an *infinitely* repeated game in extensive form
  - What are the issues?
- Infinite tree and so can not attach payoffs to terminal nodes
- Unclear how to formulate payoffs (simply summing the payoff in each stage game results in infinite payoffs)
- What are the strategies?

Infinitely Repeated Games

- Consider an *infinitely* repeated game in extensive form
  - What are the issues?
- Infinite tree and so can not attach payoffs to terminal nodes
- Unclear how to formulate payoffs (simply summing the payoff in each stage game results in infinite payoffs)
- What are the strategies?

Rewards in Infinite Games

Definition

Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for player i, the average reward of i is

$$\lim_{k\to\infty} \sum_{j=1}^{k} \frac{r_j}{k}.$$

Definition

Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for player i and a discount factor β with $0 \leq \beta \leq 1$, i's future discounted reward is

$$\sum_{j=1}^{\infty} \beta^j r_j.$$

Rewards in Infinite Games

Definition

Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for player i, the average reward of i is

$$\lim_{k\to\infty} \sum_{j=1}^{k} \frac{r_j}{k}.$$

Definition

Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for player i and a discount factor β with $0 \leq \beta \leq 1$, i's future discounted reward is

$$\sum_{j=1}^{\infty} \beta^j r_j.$$

Strategies

- A pure strategy is a choice of action at every decision point
  - That is, and action at every stage game (an infinite number of actions)
- Famous strategies for the Prisoner's Dilemma
  - Tit-for-Tat: Start out cooperating. If the opponent defects, then defect for one round. Then go back to cooperating.
  - Trigger (Grim): Start out cooperating. If the opponent defects, then defect forever.

Strategies

- A pure strategy is a choice of action at every decision point
  - That is, and action at every stage game (an infinite number of actions)
- Famous strategies for the Prisoner's Dilemma
  - Tit-for-Tat: Start out cooperating. If the opponent defects, then defect for one round. Then go back to cooperating.
  - Trigger (Grim): Start out cooperating. If the opponent defects, then defect forever.

Nash Equilibrium

- We are unable to write down an induced normal form game and then appeal to Nash's existence theorem (and it does not apply to infinite games).
- Since there are an infinite number of pure strategies, there are possibly an infinite number of equilibrium!
- It is possible to characterize the *payoffs* that are achievable under equilibrium.

Folk Theorem

- Let $v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_i, s_{-i})$
- A payoff profile $r = (r_1, \ldots, r_n)$ is *enforceable* if $\forall i \; r_i \geq v_i$.
- A payoff profile $r = (r_1, \ldots, r_n)$ is *feasible* if there exist rational, nonnegative values α_a such that for all i, we can express r_i as $\sum_{a \in A} \alpha_a u_i(a)$ with $\sigma_{a \in A} \alpha_a = 1$.
  - i.e. a convex rational combination of outcomes

Folk Theorem

Theorem

Consider any n-player game G and any payoff vector $r = (r_1, \ldots, r_n)$.

1. *If* r *is the payoff in any Nash equilibrium of the infinitely repeated G with average rewards, then for each player* i, r_i *is enforceable.*
2. *If* r *is both feasible and enforceable, then* r *is the payoff in some Nash equilibrium of the infinitely repeated G with average rewards.*

Stochastic Games

- What if we don't always repeat the same stage game?
  - Agents repeatedly play games from a *set* of normal-form games
  - The game played at any iteration depends on the previous game played and on the actions taken by all agents in that game
- A Markov Decision Process (Q, A, P, R) where
  - Q is a set of states
  - A is a set of actions
  - R is a set of rewards, associated with states
  - $P : Q \times A \times Q \mapsto [0, 1]$ is a transition probability functions such that $P(q, a, q')$ is the probability of transitioning from state q to state q' when the agent takes action a

Stochastic Games

- What if we don't always repeat the same stage game?
  - Agents repeatedly play games from a *set* of normal-form games
  - The game played at any iteration depends on the previous game played and on the actions taken by all agents in that game
- A Markov Decision Process (Q, A, P, R) where
  - Q is a set of states
  - A is a set of actions
  - R is a set of rewards, associated with states
  - $P : Q \times A \times Q \mapsto [0, 1]$ is a transition probability functions such that $P(q, a, q')$ is the probability of transitioning from state q to state q' when the agent takes action a

Stochastic Games

A stochastic game is (Q, N, A, P, R) where

- Q is a finite set of states
- $A = A_1 \times \ldots \times A_n$ where A_i is a finite set of actions available to player i
- $P : Q \times A \times Q \mapsto [0, 1]$ is the transition probability function where $P(q, a, q')$ is the probability of transitioning from state q to state q' when joint action a is played
- $R = r_1, \ldots, r_n$ where $r_i : Q \times A \mapsto \mathbb{R}$ is a real valued payoff function for player i

Strategies

- A pure strategy specifies an action conditional on every possible history
- Restricted classes of strategies
  - *Behavorial strategies*: $s_i(h_t, a_i)$ returns the probability of playing action a_i for history h_t
  - *Markov strategies:* s_i is a behavioral strategy in which $s_i(h_t, a_i) = s_i(h'_t, a_i)$ if $q_t = q'_t$ are the final states in h_t and h'_t respectively
  - *Stationary strategies:* s_i is a Markov strategy in which $s_i(h, a_i) = s_i(h', a_i)$ if $q = q'$ where q and q' are final states of h and h' respectively (i.e. no dependence on t)

Strategies

- A pure strategy specifies an action conditional on every possible history
- Restricted classes of strategies
  - *Behavorial strategies*: $s_i(h_t, a_i)$ returns the probability of playing action a_i for history h_t
  - *Markov strategies:* s_i is a behavioral strategy in which $s_i(h_t, a_i) = s_i(h'_t, a_i)$ if $q_t = q'_t$ are the final states in h_t and h'_t respectively
  - *Stationary strategies:* s_i is a Markov strategy in which $s_i(h, a_i) = s_i(h', a_i)$ if $q = q'$ where q and q' are final states of h and h' respectively (i.e. no dependence on t)

Equilibrium

- *Markov perfect equilibrium:* A strategy profile consisting only of Markov strategies that is a Nash equilibrium regardless of the starting state

Theorem

Every n-player general sum discounted reward stochastic game has a Markov perfect equilibrium.