Stochastic Games and Bayesian Games

CPSC 532l Lecture 10

Lecture Overview

Finitely Repeated Games

- Everything is straightforward if we repeat a game a finite number of times
- we can write the whole thing as an extensive-form game with imperfect information
  - at each round players don't know what the others have done; afterwards they do
  - overall payoff function is additive: sum of payoffs in stage games

Infinitely Repeated Games

- Consider an infinitely repeated game in extensive form:
  - an infinite tree!
- Thus, payoffs cannot be attached to terminal nodes, nor can they be defined as the sum of the payoffs in the stage games (which in general will be infinite).

Definition

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

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

Nash Equilibria

- With an infinite number of equilibria, what can we say about Nash equilibria?
  - we won't be able to construct an induced normal form and then appeal to Nash's theorem to say that an equilibrium exists
  - Nash's theorem only applies to finite games
- Furthermore, with an infinite number of strategies, there could be an infinite number of pure-strategy equilibria!
- It turns out we can characterize a set of payoffs that are achievable under equilibrium, without having to enumerate the equilibria.

Definitions

- Consider any n-player game $G = (N, A, u)$ and any payoff vector $r = (r_1, r_2, \ldots, r_n)$.
- Let $v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$.
  - i's minmax value: the amount of utility i can get when $-i$ play a minmax strategy against him

Definition

A payoff profile r is enforceable if $r_i \geq v_i$.

Definition

A payoff profile r is feasible if there exist rational, non-negative values α_a such that for all i, we can express r_i as $\sum_{a \in A} \alpha u_i(a)$, with $\sum_{a \in A} \alpha_a = 1$.

- a payoff profile is feasible if it is a convex, rational combination of the outcomes in G.

Folk Theorem

Theorem (Folk Theorem)

Consider any n-player game G and any payoff vector $(r_1, r_2, \ldots, r_n)$.

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.*

Folk Theorem (Part 1)

Payoff in Nash → enforceable

Part 1: Suppose r is not enforceable, i.e. $r_i < v_i$ for some i. Then consider a deviation of this player i to $b_i(s_{-i}(h))$ for any history h of the repeated game, where b_i is any best-response action in the stage game and $s_{-i}(h)$ is the equilibrium strategy of other players given the current history h. By definition of a minmax strategy, player i will receive a payoff of at least v_i in every stage game if he adopts this strategy, and so i's average reward is also at least v_i. Thus i cannot receive the payoff $r_i < v_i$ in any Nash equilibrium.

Folk Theorem (Part 2)

Feasible and enforceable → Nash

Part 2: Since r is a feasible payoff profile, we can write it as $r_i = \sum_{a \in A} \left(\frac{\beta_a}{\gamma}\right) u_i(a)$, where β_a and γ are non-negative integers.[1] Since the combination was convex, we have $\gamma = \sum_{a \in A} \beta_a$. We're going to construct a strategy profile that will cycle through all outcomes $a \in A$ of G with cycles of length γ, each cycle repeating action a exactly β_a times. Let (a^t) be such a sequence of outcomes. Let's define a strategy s_i of player i to be a trigger version of playing (a^t): if nobody deviates, then s_i plays a_i^t in period t. However, if there was a period t' in which some player $j \neq i$ deviated, then s_i will play $(p_{-j})_i$, where (p_{-j}) is a solution to the minimization problem in the definition of v_j.

[1] Recall that α_a were required to be rational. So we can take γ to be their common denominator.

Folk Theorem (Part 2)

Feasible and enforceable $\rightarrow$ Nash

First observe that if everybody plays according to s_i, then, by construction, player i receives average payoff of r_i (look at averages over periods of length γ). Second, this strategy profile is a Nash equilibrium. Suppose everybody plays according to s_i, and player j deviates at some point. Then, forever after, player j will receive his $\min\max$ payoff $v_j \leq r_j$, rendering the deviation unprofitable.

Lecture Overview

Introduction

- What if we didn't always repeat back to the same stage game?
- A stochastic game is a generalization of repeated games
  - 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 stochastic game is a generalized Markov decision process
  - there are multiple players
  - one reward function for each agent
  - the state transition function and reward functions depend on the action choices of both players

Formal Definition

Definition

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

- Q is a finite set of states,
- N is a finite set of n players,
- $A = A_1 \times \cdots \times A_n$, 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; $P(q, a, \hat{q})$ is the probability of transitioning from state s to state $\hat{q}$ after joint action a, and
- $R = r_1, \ldots, r_n$, where $r_i : Q \times A \mapsto \mathbb{R}$ is a real-valued payoff function for player i.

Remarks

- This assumes strategy space is the same in all games
  - otherwise just more notation
- Again we can have average or discounted payoffs.
- Interesting special cases:
  - zero-sum stochastic game
  - single-controller stochastic game
    - transitions (but not payoffs) depend on only one agent

Strategies

- What is a pure strategy?

Strategies

- What is a pure strategy?
  - pick an action conditional on every possible history
  - of course, mixtures over these pure strategies are possible too!
- Some interesting restricted classes of strategies:
  - behavioral strategy: $s_i(h_t, a_{i_j})$ returns the probability of playing action a_{i_j} for history h_t.
    - the substantive assumption here is that mixing takes place at each history independently, not once at the beginning of the game
  - Markov strategy: s_i is a behavioral strategy in which $s_i(h_t, a_{i_j}) = s_i(h'_t, a_{i_j})$ if $q_t = q'_t$, where q_t and q'_t are the final states of h_t and h'_t, respectively.
    - for a given time t, the distribution over actions only depends on the current state
  - stationary strategy: s_i is a Markov strategy in which $s_i(h_{t_1}, a_{i_j}) = s_i(h'_{t_2}, a_{i_j})$ if $q_{t_1} = q'_{t_2}$, where q_{t_1} and q'_{t_2} are the final states of h_{t_1} and h'_{t_2}, respectively.
    - no dependence even on t

Equilibrium (discounted rewards)

- Markov perfect equilibrium:
  - a strategy profile consisting of only Markov strategies that is a Nash equilibrium regardless of the starting state
  - analogous to subgame-perfect equilibrium

Theorem

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

Equilibrium (average rewards)

- Irreducible stochastic game:
  - every strategy profile gives rise to an irreducible Markov chain over the set of games
    - irreducible Markov chain: possible to get from every state to every other state
  - during the (infinite) execution of the stochastic game, each stage game is guaranteed to be played infinitely often—for any strategy profile
  - without this condition, limit of the mean payoffs may not be defined

Theorem

For every 2-player, general sum, average reward, irreducible stochastic game has a Nash equilibrium.

A folk theorem

Theorem

For every 2-player, general sum, irreducible stochastic game, and every feasible outcome with a payoff vector r that provides to each player at least his minmax value, there exists a Nash equilibrium with a payoff vector r. This is true for games with average rewards, as well as games with large enough discount factors (i.e. with players that are sufficiently patient).

Lecture Overview

Fun Game

- Choose a phone number none of your neighbours knows; consider it to be ABC-DEFG

Fun Game

- Choose a phone number none of your neighbours knows; consider it to be ABC-DEFG
  - take "DE" as your valuation
  - play a first-price auction with three neighbours, where your utility is your valuation minus the amount you pay

Fun Game

- Choose a phone number none of your neighbours knows; consider it to be ABC-DEFG
  - take “DE” as your valuation
  - play a first-price auction with three neighbours, where your utility is your valuation minus the amount you pay
  - now play the auction again, same neighbours, same valuation

Fun Game

- Choose a phone number none of your neighbours knows; consider it to be ABC-DEFG
  - take "DE" as your valuation
  - play a first-price auction with three neighbours, where your utility is your valuation minus the amount you pay
  - now play the auction again, same neighbours, same valuation
  - now play again, with "FG" as your valuation

Fun Game

- Choose a phone number none of your neighbours knows; consider it to be ABC-DEFG
  - take "DE" as your valuation
  - play a first-price auction with three neighbours, where your utility is your valuation minus the amount you pay
  - now play the auction again, same neighbours, same valuation
  - now play again, with "FG" as your valuation
- Questions:
  - what is the role of uncertainty here?
  - can we model this uncertainty using an imperfect information extensive form game?

Fun Game

- Choose a phone number none of your neighbours knows; consider it to be ABC-DEFG
  - take "DE" as your valuation
  - play a first-price auction with three neighbours, where your utility is your valuation minus the amount you pay
  - now play the auction again, same neighbours, same valuation
  - now play again, with "FG" as your valuation
- Questions:
  - what is the role of uncertainty here?
  - can we model this uncertainty using an imperfect information extensive form game?
    - imperfect info means not knowing what node you're in in the info set
    - here we're not sure what game is being played (though if we allow a move by nature, we can do it)

Introduction

- So far, we've assumed that all players know what game is being played. Everyone knows:
  - the number of players
  - the actions available to each player
  - the payoff associated with each action vector
- Why is this true in imperfect information games?
- We'll assume:

1. All possible games have the same number of agents and the same strategy space for each agent; they differ only in their payoffs.
2. The beliefs of the different agents are posteriors, obtained by conditioning a common prior on individual private signals.

Definition 1: Information Sets

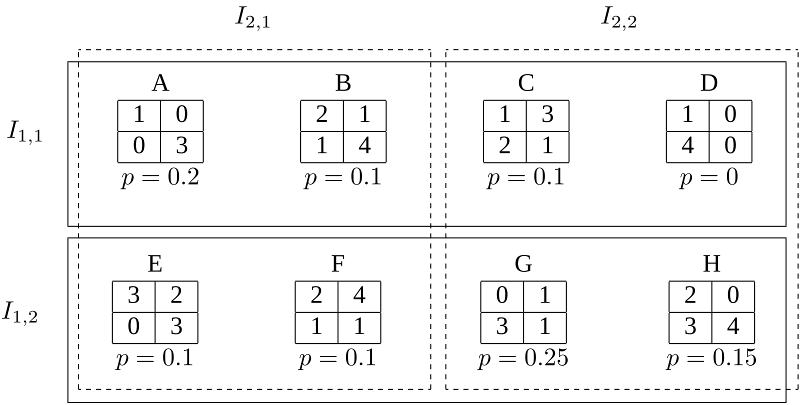
- Bayesian game: a set of games that differ only in their payoffs, a common prior defined over them, and a partition structure over the games for each agent.

Definition (Bayesian Game: Information Sets)

A Bayesian game is a tuple (N, G, P, I) where

- N is a set of agents,
- G is a set of games with N agents each such that if $g, g' \in G$ then for each agent $i \in N$ the strategy space in g is identical to the strategy space in g',
- $P \in \Pi(G)$ is a common prior over games, where $\Pi(G)$ is the set of all probability distributions over G, and
- $I = (I_1, ..., I_N)$ is a set of partitions of G, one for each agent.

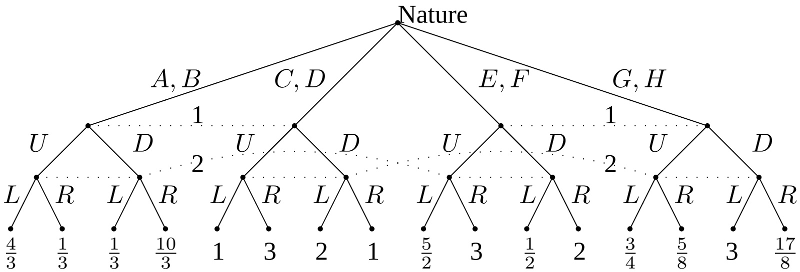
Definition 1: Example



Definition 2: Extensive Form with Chance Moves

- Add an agent, "Nature," who follows a commonly known mixed strategy.
- Thus, reduce Bayesian games to extensive form games of imperfect information.
- This definition is cumbersome for the same reason that IIEF is a cumbersome way of representing matrix games like Prisoner's dilemma
  - however, it makes sense when the agents really do move sequentially, and at least occasionally observe each other's actions.

Definition 2: Example



Definition 3: Epistemic Types

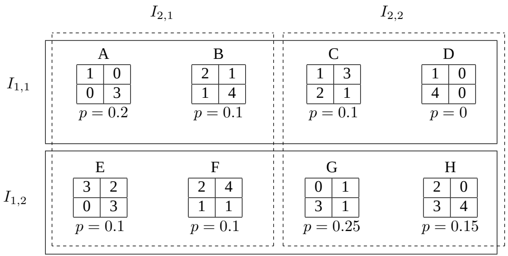
- Directly represent uncertainty over utility function using the notion of epistemic type.

Definition

A Bayesian game is a tuple (N, A, Θ, p, u) where

- N is a set of agents,
- $A = (A_1, \ldots, A_n)$, where A_i is the set of actions available to player i,
- $\Theta = (\Theta_1, \ldots, \Theta_n)$, where Θ_i is the type space of player i,
- $p : \Theta \rightarrow [0, 1]$ is the common prior over types,
- $u = (u_1, \ldots, u_n)$, where $u_i : A \times \Theta \rightarrow \mathbb{R}$ is the utility function for player i.

Definition 3: Example



a_1	a_2	θ_1	θ_2	u_1
U	L	$\theta_{1,1}$	$\theta_{2,1}$	4/3
U	L	$\theta_{1,1}$	$\theta_{2,2}$	1
U	L	$\theta_{1,2}$	$\theta_{2,1}$	5/2
U	L	$\theta_{1,2}$	$\theta_{2,2}$	3/4
U	R	$\theta_{1,1}$	$\theta_{2,1}$	1/3
U	R	$\theta_{1,1}$	$\theta_{2,2}$	3
U	R	$\theta_{1,2}$	$\theta_{2,1}$	3
U	R	$\theta_{1,2}$	$\theta_{2,2}$	5/8

a_1	a_2	θ_1	θ_2	u_1
D	L	$\theta_{1,1}$	$\theta_{2,1}$	1/3
D	L	$\theta_{1,1}$	$\theta_{2,2}$	2
D	L	$\theta_{1,2}$	$\theta_{2,1}$	1/2
D	L	$\theta_{1,2}$	$\theta_{2,2}$	3
D	R	$\theta_{1,1}$	$\theta_{2,1}$	10/3
D	R	$\theta_{1,1}$	$\theta_{2,2}$	1
D	R	$\theta_{1,2}$	$\theta_{2,1}$	2
D	R	$\theta_{1,2}$	$\theta_{2,2}$	17/8

Lecture Overview

1. Recap
2. Stochastic Games
3. Bayesian Games
4. **Analyzing Bayesian games**

Strategies

- Pure strategy: $s_i : \Theta_i \rightarrow A_i$
  - a mapping from every type agent i could have to the action he would play if he had that type.
- Mixed strategy: $s_i : \Theta_i \rightarrow \Pi(A_i)$
  - a mapping from i's type to a probability distribution over his action choices.
- $s_j(a_j|\theta_j)$
  - the probability under mixed strategy s_j that agent j plays action a_j, given that j's type is θ_j.

Expected Utility

Three meaningful notions of expected utility:

- *ex-ante*
  - the agent knows nothing about anyone's actual type;
- *ex-interim*
  - an agent knows his own type but not the types of the other agents;
- *ex-post*
  - the agent knows all agents' types.

Ex-interim expected utility

Definition (*Ex-interim* expected utility)

Agent i's *ex-interim* expected utility in a Bayesian game (N, A, Θ, p, u), where i's type is θ_i and where the agents' strategies are given by the mixed strategy profile s, is defined as

$$EU_i(s|\theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i}|\theta_i) \sum_{a \in A} \left(\prod_{j \in N} s_j(a_j|\theta_j) \right) u_i(a, \theta_{-i}, \theta_i).$$

- i must consider every θ_{-i} and every a in order to evaluate $u_i(a, \theta_i, \theta_{-i})$.
- i must weight this utility value by:
  - the probability that a would be realized given all players' mixed strategies and types;
  - the probability that the other players' types would be θ_{-i} given that his own type is θ_i.

Ex-ante expected utility

Definition (*Ex-ante* expected utility)

Agent i's *ex-ante* expected utility in a Bayesian game (N, A, Θ, p, u), where the agents' strategies are given by the mixed strategy profile s, is defined as

$$EU_i(s) = \sum_{\theta_i \in \Theta_i} p(\theta_i) EU_i(s|\theta_i)$$

or equivalently as

$$EU_i(s) = \sum_{\theta \in \Theta} p(\theta) \sum_{a \in A} \left(\prod_{j \in N} s_j(a_j|\theta_j) \right) u_i(a, \theta).$$

Ex-post expected utility

Definition (*Ex-post* expected utility)

Agent i's *ex-post* expected utility in a Bayesian game (N, A, Θ, p, u), where the agents' strategies are given by s and the agent' types are given by θ, is defined as

$$EU_i(s,\theta) = \sum_{a \in A} \left(\prod_{j \in N} s_j(a_j|\theta_j) \right) u_i(a,\theta).$$

- The only uncertainty here concerns the other agents' mixed strategies, since i knows everyone's type.

Best response

Definition (Best response in a Bayesian game)

The set of agent i's best responses to mixed strategy profile s_{-i} are given by

$$BR_i(s_{-i}) = \arg\max_{s'_i \in S_i} EU_i(s'_i, s_{-i}).$$

- it may seem odd that BR is calculated based on i's *ex-ante* expected utility.
- However, write $EU_i(s)$ as $\sum_{\theta_i \in \Theta_i} p(\theta_i) EU_i(s|\theta_i)$ and observe that $EU_i(s'_i, s_{-i}|\theta_i)$ does not depend on strategies that i would play if his type were not θ_i.
- Thus, we are in fact performing independent maximization of i's *ex-interim* expected utility conditioned on each type that he could have.

Nash equilibrium

Definition (Bayes-Nash equilibrium)

A Bayes-Nash equilibrium is a mixed strategy profile s that satisfies $\forall i \;\; s_i \in BR_i(s_{-i})$.

- we can also construct an induced normal form for Bayesian games
- the numbers in the cells will correspond to *ex-ante* expected utilities
  - however as argued above, as long as the strategy space is unchanged, best responses don't change between the *ex-ante* and *ex-interim* cases.

ex-post Equilibrium

Definition (*ex-post* equilibrium)

A *ex-post* Bayes-Nash equilibrium is a mixed strategy profile s that satisfies $\forall\theta,\ \forall i,\ s_i \in \arg\max_{s'_i \in S_i} EU_i(s'_i, s_{-i}, \theta)$.

- somewhat similar to dominant strategy, but not quite
  - EP: agents do not need to have accurate beliefs about the type distribution
  - DS: agents do not need to have accurate beliefs about others' strategies