Game-theoretic
Foundations of Multi-agent Systems

Lecture 3: Games in Extensive Form

Seyed Majid Zahedi

UNIVERSITY OF
WATERLOO

Outline

Extensive-form Games

- So far, we have studied strategic-form games
  - Agents take actions once and simultaneously
- Next, we study extensive-form games (a.k.a. sequential or multi-stage games)
  - Extensive-form games can be conveniently represented by game trees

(Finite) Perfect-info Extensive-form Game: Definition

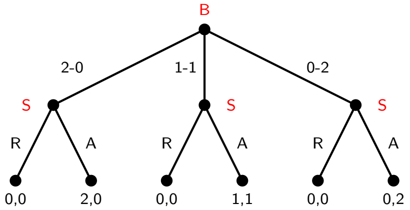
- The game consists of a set of agents, $N = \{1, 2, \ldots, n\}$
- A is set of actions
- H is set of choice nodes (internal nodes of game tree)
- Z is set of terminal nodes (leaves of game tree)

(Finite) Perfect-info Extensive-form Game: Definition (cont.)

- $\alpha : H \rightarrow N$ is agent function
  - Maps each choice node to an agent who chooses an action at that node
- $\beta : H \rightarrow 2^A$ is action function
  - Maps each choice node to set of actions available at that node
- $\rho : H \times A \rightarrow H \cup Z$ is successor function
  - Maps each choice node and action pair to new choice node or terminal node
  - If $\rho(h_1, a_1) = \rho(h_2, a_2)$ then $h_1 = h_2$ and $a_1 = a_2$
- $u = (u_1, \ldots, u_n)$, where $u_i : Z \rightarrow \mathbb{R}$ is agent i's utility function
  - Maps each terminal node to a real value

Example: Sharing Game

- Brother and sister share two gifts
- Brother suggests a split first
- Sister then chooses to accept or reject
- If she accepts, they get suggested gifts
- Otherwise, neither gets any gift



Outline

History in Extensive-form Games

- If height of game tree (i.e, number of stages) is finite, then game is finite-horizon game
- Otherwise, the game is called infinite-horizon game
- For perfect-information games, each node maps to unique history (and vice versa)
- Since choice nodes form a tree, we can unambiguously identify a node with its history
  - I.e., sequence of choices leading from the root node to it

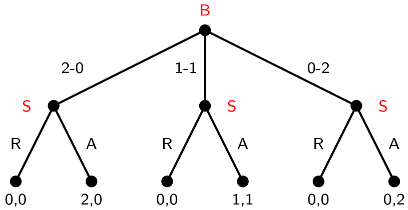
Pure Strategies

- Agent i's pure strategy defines contingency plan for <u>all</u> choice nodes mapped to i

$$a_i \in A_i = \prod_{h \in H, \alpha(h) = i} \beta(h)$$

- Strategy must specify a decision at each choice node
  - Regardless of whether it is possible to reach that node

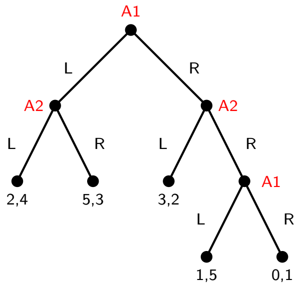
Pure Strategies: Example



- $A_B = \{$"2-0", "1-1", "0-2"$\}$
- $A_S = \{$(R, R, R), (R, R, A), (R, A, R), (A, R, R), (R, A, A), (A, R, A), (A, A, R), (A, A, A)$\}$

Pure Strategies: (Another) Example

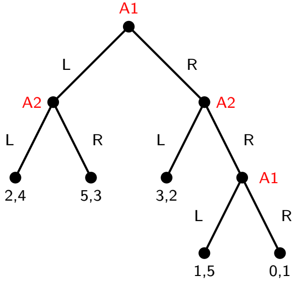
- What are pure strategies for A2?
  - $A_{A2} = \{(\text{L, L}), (\text{L, R}), (\text{R, L}), (\text{R, R})\}$
- What about A1?
  - $A_{A1} = \{(\text{L, L}), (\text{L, R}), (\text{R, L}), (\text{R, R})\}$



Normal-form Representation of Extensive-form Games

- For every perfect-info game, there is corresponding normal-form game

A1 \ A2	(L, L)	(L, R)	(R, L)	(R, R)
(L, L)	2, 4	2, 4	5, 3	5, 3
(L, R)	2, 4	2, 4	5, 3	5, 3
(R, L)	3, 2	1, 5	3, 2	1, 5
(R, R)	3, 2	0, 1	3, 2	0, 1



Transformation from Extensive form to Normal From

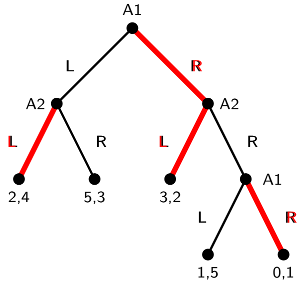
- It can always be performed for perfect-information games
- It can cause redundancy
  - E.g., $(2, 4)$ occurs once in extensive form but 4 times in normal form
- It can result in exponential blowup of game representation
- Reverse transformation does not always exist
  - E.g., there is no extensive-form representation for Prisoner's Dilemma
  - Perfect-information extensive-form games cannot model simultaneity

Nash Equilibrium of Perfect-info Games in Extensive Form

- [Theorem] Every (finite) perfect-info extensive-form game has pure-strategy NE
- Agents see everything before each action $\Rightarrow$ randomness is not required
- This is not the case for every finite game in normal form

Nash Equilibrium: An Empty Threat?

A1 \ A2	(L, L)	(L, R)	(R, L)	(R, R)
(L, L)	2, 4	2, 4	5, 3	5, 3
(L, R)	2, 4	2, 4	5, 3	5, 3
(R, L)	3, 2	1, 5	3, 2	1, 5
(R, R)	3, 2	0, 1	3, 2	0, 1



- Strategy of A1 is called a threat
  - Committing to choose R forces A2 to avoid that part of the tree
- A2 may not consider A1's threat to be credible
  - Would A1 really follow through on this threat if final decision node is reached?

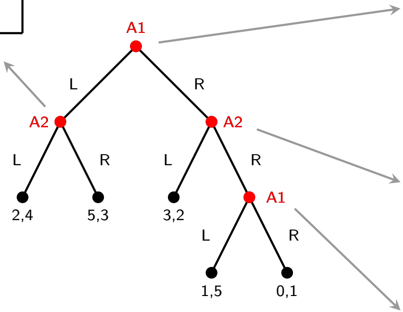
Outline

Subgames: Definition

- Let G be a perfect-information extensive-form game
- Subgame of G rooted at node h is restriction of G to descendants of h
- Set of subgames of G consists of all of subgames of G rooted at some node in G

Subgames: Example

		A2	
		(*, L)	(*, R)
A1	(*, *)	2,4	5,3



		A2			
		(L, L)	(L, R)	(R, L)	(R, R)
A1	(L, L)	2,4	2,4	5,3	5,3
	(L, R)	2,4	2,4	5,3	5,3
	(R, L)	3,2	1,5	3,2	1,5
	(R, R)	3,2	0,1	3,2	0,1

		A2	
		(*, L)	(*, R)
A1	(*, L)	3,2	1,5
	(*, R)	3,2	0,1

		A2
		(*, *)
A1	(*, L)	1,5
	(*, R)	0,1

Subgame Perfect Equilibrium (SPE)

- Let $s_{G'}$ be restriction of strategy profile s to subgame G'
- Profile s^* is SPE of game G if for every subgame G' of $G, s^*_{G'}$ is NE
- Loosely speaking, subgame perfection removes non-credible threats
  - Non-credible threads are not NE in their subgames
- How to find SPE?
  - One could find all of NE, then eliminate those that are not subgame perfect
  - But there are more economical ways of doing it

Computing Equilibrium: Backward Induction for Finite Games

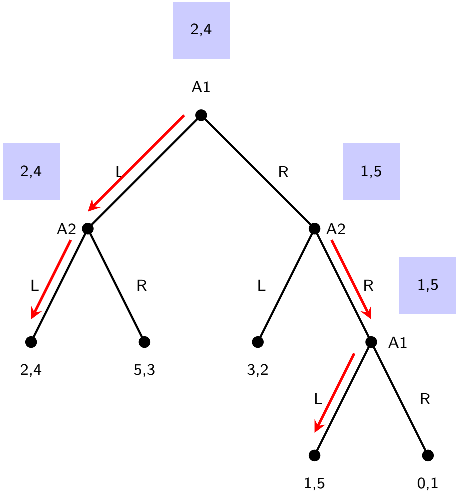
- (1) Start from "last" subgames (choice nodes with all terminal children)
- (2) Find Nash equilibria of those subgames
- (3) Turn those choice nodes to terminal nodes using NE utilities
- (4) Go to (1) until no choice node remains

Backward Induction Procedure

Algorithm 1: Finding value of sample SPE of perfect-info extensive-form game

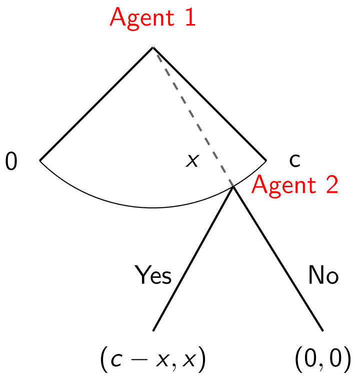
procedure `Backward_Induction`(*node h*)
- **if** $h \in Z$ **then**
  - **return** $u(h)$;
- *best_utility* $\leftarrow -\infty$;
- **forall** $a \in \beta(h)$ **do**
  - $u =$ `Backward_Induction`$(\rho(h, a))$;
  - **if** $u_{\alpha(h)} >$ *best_utility* **then**
    - *best_utility* $= u_{\alpha(h)}$;
- **return** *best_utility*

SPE: Example



Example: Ultimatum Game

- Two agents want to split c dollars
  - A1 offers A2 some amount $x \leq c$
  - If A2 accepts, outcome is $(c - x, x)$
  - If A2 rejects, outcome is $(0, 0)$
- What is A2's best response if $x > 0$?
  - Yes
- What is A2's best response if $x = 0$?
  - Indifferent between Yes or No
- What are A2's optimal strategies?
  - Option 1: Yes for all $x \geq 0$
  - Option 2: Yes if $x > 0$, No if $x = 0$



SPE of Ultimatum Game

- What is A1's optimal strategy for each of A2's optimal strategies?
  - For option 1, A1's optimal strategy is to offer $x = 0$
  - For option 2, if A1 offers $x = 0$, then A1's utility is 0
  - If A1 wants to offer any $x > 0$, then A1 must offer

$$\operatorname*{argmax}_{x>0}(c - x)$$

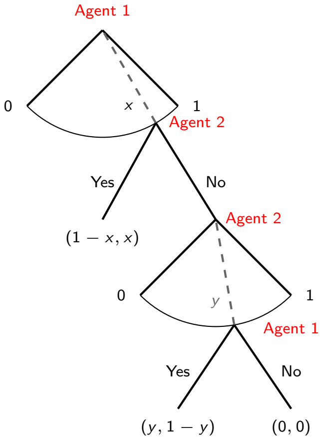
  - This optimization does not have any optimal solution
  - No offer of agent 1 is optimal
- Unique SPE of ultimatum game is A1 offers 0, and A2 accepts all offers

Example: Discrete Ultimatum Game

- What are A2's optimal strategies if c is in multiple of cent?
  - Option 1: Yes for all $x \geq 0$
  - Option 2: Yes if $x > 0$, No if $x = 0$
- What are A1's optimal strategies for each of A2's?
  - For option 1, offer $x = 0$
  - For option 2, offer $x = 1$ cent
- What are SPE of this modified ultimatum game?
  - A1 offers 0, and A2 accepts all offers
  - A1 offers 1 cent, and A2 accepts all offers except 0
- Show that every $\bar{x} \in [0, c]$, there exists NE in which A1 offers $\bar{x}$
  - What is agent A2's optimal strategy?

Example: Bargaining Game

- Two agents want to split $c = 1$ dollar
- First, A1 makes her offer
- Then, A2 decides to accept or reject
- If A2 rejects, then A2 makes new offer
- Then, A1 decides to accept or reject
- Let $x = (x_1, x_2)$ denote A1's offer
- Let $y = (y_1, y_2)$ denote A2's offer

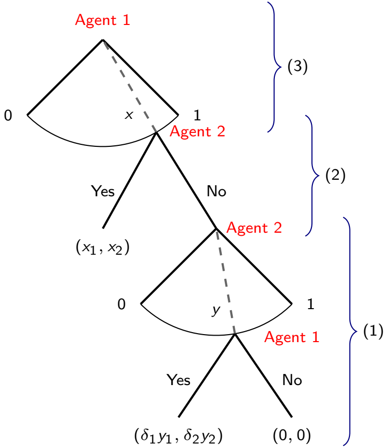


Backward Induction for Bargaining Game

- Second round is ultimatum game with unique SPE
  - A2 offers 0, and A1 accepts all offers
- What is A2's optimal strategy in round 1's subgame?
  - Option 1: If $x_2 \leq 1$, reject
  - Option 2: If $x_2 = 1$, accept, and reject otherwise
- What are A1's optimal strategies in round 1 for each of A2's?
  - For both options, A1 is indifferent between all strategies
  - A1's weakly dominant strategy is to offer $x_2 = 1$
- How many SPE does this game have?
  - Infinitely many! In all SPE, A2 gets everything (Last mover's advantage)
  - In every SPE, agent who makes offer in last round gets everything

Example: Discounted Bargaining Game

- Utilities are discounted by $0 < \delta_i < 1$
- What is unique SPE of (1)?
  - A2 offers $y_1 = 0$ and A1 accepts all offers
- What are optimal strategies in (2)?
  - Option 1: Yes if $x_2 \geq \delta_2$, No otherwise
  - Option 2: Yes if $x_2 > \delta_2$, No otherwise
- What are optimal strategies in (3)?
  - For option 1, offer $x_2 = \delta_2$
  - For option 2, there is no optimal strategy



Unique SPE of Discounted Bargaining Game

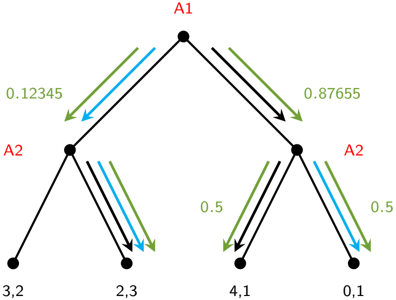
- What are SPE strategies?
  - Agent 1's proposes $(1-\delta_2, \delta_2)$
  - Agent 2 only accepts proposals with $x_2 \geq \delta_2$
  - Agent 2 proposes $(0, 1)$ after any history in which1's proposal is rejected
  - Agent 1 accepts all proposals of Agent 2

- What is SPE outcome of game?
  - Agent 1 proposes $(1-\delta_2, \delta_2)$
  - Agent 2 accepts
  - Resulting utilities are $(1-\delta_2, \delta_2)$

- Desirability of earlier agreement yields positive utility for agent 1

Limitation of Backward Induction

- If there are ties, how they are broken affects what happens up in tree



Outline

Imperfect-info Games: Motivation

- So far, we have allowed agents to specify action they take at every choice node
- This implies that agents know the node they are in and all prior choices
- This is why we call these games perfect-information games
- However, this might not be the case in all environments

Imperfect-info Games: Motivation (cont.)

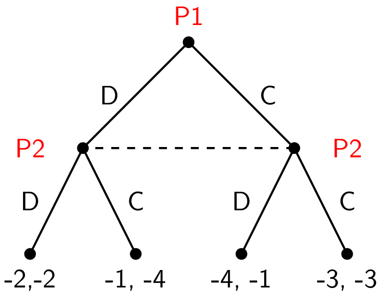
- We may want to model agents with partial or no knowledge of others' actions
- We may even want to model agents with limited memory of their own past actions
- Imperfect-info games in extensive form address this limitation
- In such games, each agent's choice nodes are partitioned into information sets
- If two nodes are in same info set, then agent cannot distinguish between them

Imperfect-info Extensive-form Games: Definition

- N, A, H, Z, α, β, ρ, u are the same as before
- $I = (I_1, ..., I_n)$, where $I_i = (I_{i,1}, ..., I_{i,k_i})$ is a partition of $\{h \in H : \alpha(h) = i\}$
- If h, h' are in the same equivalence class $I_{i,j}$, then $\beta(h) = \beta(h')$
- Perfect-info games are imperfect-info games with singleton equivalence classes

Example: Prisoners' Dilemma in Extensive Form

- P1 decides on D or C
- P2 then decides on D or C (without observing P1's decision)



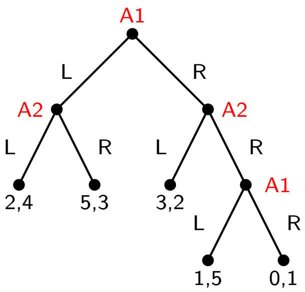
Outline

Pure, Mixed, and Behavioral Strategies

- Pure strategies of agent i consists of $\prod_{I_{i,j} \in I_i} \beta(I_{i,j})$
- Mixed strategies define randomization over pure strategies
- Behavioral strategy define independent randomization at each info set
- Mixed strategy is distribution over vectors (each vector describing a pure strategy)
- Behavioral strategy is a vector of distributions
- In general, expressive power of behavioral and mixed strategies are noncomparable
  - In some games, there are outcomes that are achieved via mixed but not any behavioral strategies
  - And in some games it is the other way around

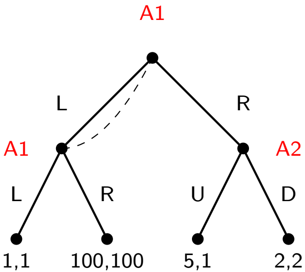
Mixed vs Behavioral Strategies: Example I

- Give behavioral strategy for A1
  - L w.p. 0.2 and L w.p. 0.5
- Give mixed strategy for A1 that is not behavioral strategy
  - (L, L) w.p. 0.4 and (R, R) w.p. 0.6
  - Why this is not behavioral strategy?
- In this game, every behavioral strategy corresponds to a mixed strategy and vice versa (more on this soon)



Mixed vs Behavioral Strategies: Example II

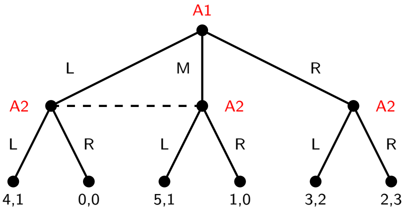
- What is mixed-strategy NE of this game?
  - (R, D) with outcome utilities (2,2)
- What is A1's expected utility for $(p, 1-p)$?
  - $p^2 + 100p(1-p) + 2(1-p)$
- What is A1's best response?
  - $p = 98/198$
- What is behavioral NE of this game?
  - ((98/198, 100/198), (0, 1))



Perfect Recall

- Strategies that induce same distribution on outcomes, for fixed strategy profile of others, are called equivalent strategies
- If all agents remember all their own actions, game is a game of perfect recall
- In such games, any mixed strategy of given agent can be replaced by an equivalent behavioral strategy
- And any behavioral strategy can be replaced by an equivalent mixed strategy

Subgame Perfection and Imperfect Information



- There are two subgames: game itself and subgame after agent 1 plays R
  - (R, (R,R)) is NE and SPE
- But, why should 2 play R after 1 plays L or M?
  - This is non-credible threat
- There are more sophisticated equilibrium refinements that rule this out
  - They explicitly model agents' beliefs on where they are for every info set
  - E.g., sequential equilibrium, perfect Bayesian equilibrium

Acknowledgment

- This lecture is a slightly modified version of ones prepared by
  - Asu Ozdaglar [MIT 6.254]
  - Vincent Conitzer [Duke CPS 590.4]
- Hadi Omidi helped with importing slides from PowerPoint to LaTeX