NUS National University of Singapore | Computing

Game Theory

CS4246/CS5446

AI Planning and Decision Making

Please join:
pollev.com/anarayan

This lecture will be recorded!

Topics

- Non cooperative game theory (17.2)
  - Single Move Games - Normal Form (17.2.1)
  - Computing Equilibria (17.2.2)

Recall: Decision Theory

- Utility theory:
  - Quantifies degrees of preferences of alternatives
- Decision Model:
  - Explicates impact of uncertainty on these preferences
- Utility function:
  - Mapping from states of the world to real numbers, indicating agent's level of happiness with that state of the world
- Decision-theoretic rationality:
  - Takes actions to maximize expected utility.

Game Theory: Overview

- Multiagent decision making
  - Simultaneous moves and partial observability
  - Perfect information vs imperfect information
- Agent design
  - Analyze agent's decisions; compute expected utility for each decision for optimal agents
  - E.g., Two finger Morra
- Mechanism design
  - Define rules for the environment with multiple agents to maximize collective good
  - E.g., Protocols for internet traffic routers, intelligent multiagent systems
- Assumption:
  - All agents obey optimal decision making rules according to game theory

Elements of A Game

- Rules of the Game: PAPI
  - Players
  - Actions
  - Payoffs
  - Information
- Objective
  - Describe a situation in terms of the rules of the game
  - Players will maximize payoffs
  - Players will devise strategies to pick actions depending on information available
  - Combination of strategies chosen by each player is the equilibrium
  - Given an equilibrium, predict outcome of the game

Quiz

Quiz

Quiz

Playing Games

- Which of the following is a game?
  1. OPEC members choosing their annual output
  2. Company deciding to hire 10 senior programmers
  3. Apple purchasing OLED displays from Samsung

Source: Adapted from Rasmussen 2005, Chapter 1

Non-Cooperative Games

- What are non-cooperative games?
  - Mathematical study of interaction between rational, self-interested agents
- Why is it called non-cooperative?
  - While most interested in situations where agents’ interests conflict, not restricted to these settings
  - Key is that the agent is the basic modeling unit (with beliefs, preferences, possible actions), and that agents pursue their own interests
- What is a self-interested agent?
  - Not that it wants to harm other agents
  - Not that it cares about only things that benefit it
  - Agent has a description of states of the world that it likes; its actions are motivated by this description

Single-Move Games

Normal form games
Pure strategy games

Single-Move Games

- Definition:
  - All players take actions “simultaneously”
  - Result of game is based on this single set of actions
- Examples:
  - Prisoner’s Dilemma; Two finger Morra; etc.
  - Many real decision making situations in business, politics, defense, financial planning, etc.
- A single-move game is defined by 3 components
  1. Players or agents who make the decision
     - Focus on 2-player games; n-player games for $n > 2$ are also common
  2. Actions that players can choose
  3. Payoff function that gives the utility to each player for each combination of actions by all players
     - Can be represented as a payoff matrix – called the strategic form or normal form
     - Row Player: Player 1; Column Player: Player 2

Example: Prisoner's Dilemma

	$Al: Testify$	$Al: Refuse$
$Bo: Testify$	$B = -5; A = -5$	$B = 0; A = -10$
$Bo: Refuse$	$B = -10; A = 0$	$B = -1; A = -1$

- Game:
  - Two suspects: Bo and Al, are caught near a burglary scene and interrogated separately
  - Choices:
    - One Testifies (blames the other) and the other Refuses (to cooperate): Payoffs – [0, -10] or [-10, 0]
    - Both Refuse: Payoffs – [-1, -1]
    - Both Testify: Payoffs – [-5, -5]
  - Should each of them testify?

Example: Prisoner's Dilemma

- What should Al and Bo do?

	$Al: Testify$	$Al: Refuse$
$Bo: Testify$	$\boxed{B = -5; A = -5}$	$B = 0; A = -10$
$Bo: Refuse$	$B = -10; A = 0$	$B = -1; A = -1$

Both should testify!

General Form of Single-Move Games

- Prisoner's dilemma written as a matrix
  - "strategic form" or "normal form"

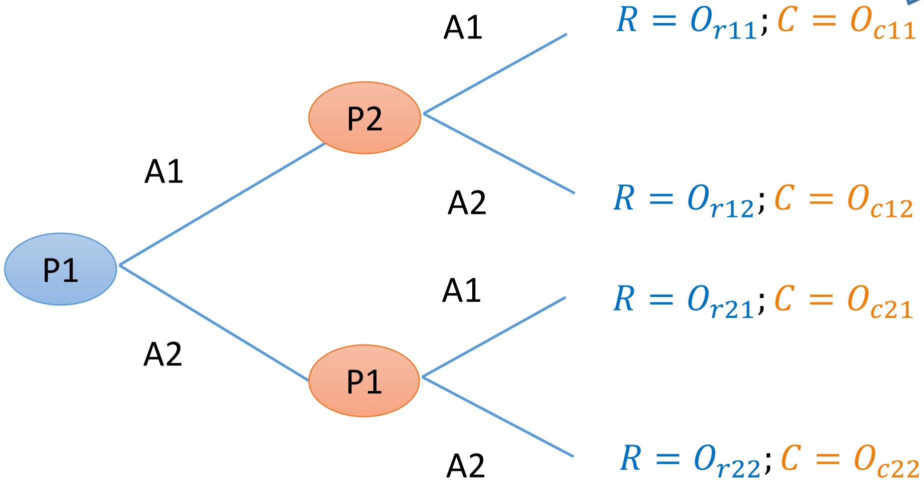
Subscripts in payoffs denote action numbers of P1 and P2

	Column Player (P2) Action 1	Column Player (P2) Action 2
Row Player (P1) Action 1	$R = O_{r11}; C = O_{c11}$	$R = O_{r12}; C = O_{c12}$
Row Player (P1) Action 2	$R = O_{r21}; C = O_{c21}$	$R = O_{r22}; C = O_{c22}$

General Form of Single-Move Games

- Prisoner's dilemma written as a game tree
  - "extensive form"

Subscripts in payoffs denote action numbers of P1 and P2



A1: Action 1
A2: Action 2

Game Strategies

- Strategy
  - Each player must adopt and execute a strategy (policy)
- Pure strategy
  - Deterministic policy
  - A single action for a single-move game
- Mixed strategy
  - Randomized policy that selects actions according to a probability distribution
  - [p: a; (1 – p): b]
  - chooses action a with probability p and b otherwise
- Strategy profile
  - Assignment of a strategy to each player

Game Solutions

- Outcome
  - Given the strategy profile, the game's outcome is a numeric value for each player (number or expected utility)
- Solution
  - A strategy profile in which each player adopts a rational strategy
  - Examples: Pareto optima and Nash equilibrium
- What is "rational"?
  - Each agent chooses only part of the strategy profile that determines the outcome
  - Outcomes are actual results of playing a game
  - Solutions are theoretical constructs for analyzing a game

Example: Prisoner's Dilemma

- What should Al and Bo do?

For both Al and Bo, $Testify$ is a dominant strategy

Dominant Strategy Equilibrium

	$Al: Testify$	$Al: Refuse$
$Bo: Testify$	$\boxed{B = -5; A = -5}$	$B = 0; A = -10$
$Bo: Refuse$	$B = -10; A = 0$	$B = -1; A = -1$

Both should testify!

Domination

- Dominant strategy
  - A strategy that dominates all others
  - Irrational to play a dominated strategy
  - Irrational not to play a dominant strategy if one exists
- Strongly dominant strategy
  - Strategy s for player P strongly dominates strategy s’ if the outcome for s is better for P than the outcome for s’, for every choice of strategies by other player(s)
- Weakly dominant strategy
  - Strategy s weakly dominates s’ if s is better than s’ on at least one strategy profile and no worse on any other

Deriving Solution with Dominated Strategies

- One way to solve the game: iteratively eliminate dominated strategies
  - If we end up with only one strategy for each player, we have a solution!
- Al first reasons that $Refuse$ is dominated, so it can be eliminated
- Simplified game after eliminating Al's $Refuse$ strategy
- In the simplified game, Bo reasons that $Refuse$ is dominated
- On eliminating, we get the solution!

	$Al: Testify$	$Al: Refuse$
$Bo: Testify$	$B = -5; A = -5$	$B = 0; A = -10$
$Bo: Refuse$	$B = -10; A = 0$	$B = -1; A = -1$

	$Al: Testify$
$Bo: Testify$	$B = -5; A = -5$
$Bo: Refuse$	$B = -10; A = 0$

	$Al: Testify$
$Bo: Testify$	$\boxed{B = -5; A = -5}$

Dominant Strategy Equilibrium

Deriving Solution with Dominated Strategies

- Strongly dominant strategy
  - If iterated elimination of strictly dominated strategies result in one strategy for each player, we have found <u>a unique equilibrium</u>
  - Order of the elimination does not matter
- Weakly dominant strategy
  - If we also eliminated weakly dominated strategies and end up with one strategy for each player, we have also found an equilibrium, but there may be other equilibriums that we did not find

Equilibrium

E.g., in Prisoner's Dilemma, both players testifying

- Dominant strategy equilibrium
  - When each player has a dominant strategy, combination of strategies is called a dominant strategy equilibrium
  - It is an "equilibrium" because no player has any incentive to deviate from their part of it: by definition, if they did so, they could not do better, and may do worse.
- Nash Equilibrium
  - A strategy profile forms an equilibrium if no other player can benefit by switching strategies, given that every other player sticks with the same strategy
  - Local optimum in the space of strategies
- Note:
  - Every game has at least one equilibrium (John Nash)
  - A dominant strategy equilibrium is a Nash equilibrium
  - Some games have Nash equilibrium but no dominant strategies

Pareto Optimality

- An outcome is Pareto dominated by another outcome if all players would prefer the other outcome
- An outcome is Pareto optimal if there is no other outcomes that all players would prefer, i.e., if there is no other outcome that would make one player better off without making someone else worse off
  - An outcome is Pareto-optimal if there is no other outcome that Pareto-dominates it.
  - A game can have more than one Pareto-optimal outcome
  - Every game has at least one Pareto-optimal outcome

Note: Pareto optimality is named for the Italian economist Vilfredo Pareto (1848–1923).

Example: Prisoner's Dilemma

Equilibrium

Pareto optimal

Pareto optimal

Pareto dominating

	$Al: Testify$	$Al: Refuse$
$Bo: Testify$	$B = -5; A = -5$	$B = 0; A = -10$
$Bo: Refuse$	$B = -10; A = 0$	$B = -1; A = -1$

- Dilemma:
  - Equilibrium outcome of both Al & Bo testifying is worse than the outcome both would get if they refuse
  - The $(Testify, Testify)$ equilibrium is pareto dominated by $(Refuse, Refuse)$ strategy profile which is better for both players
  - Although$(Refuse, Refuse)$ combination is better, both players playing rationally end up in $(Testify, Testify)$ solution

Quiz

Quiz

Quiz

Example: Prisoner's Dilemma

	$Al: Testify$	$Al: Refuse$
$Bo: Testify$	$B = -5; A = -5$	$B = 0; A = -10$
$Bo: Refuse$	$B = -10; A = 0$	$B = -1; A = -1$

- By modifying the game, it may be possible to end up with $(Refuse, Refuse)$ solution.
- For example:
  - In a repeated game, the players will meet again and again
  - Agents may have moral beliefs that encourage cooperation and fairness, changing the payoff matrix

Example: Finding Pure Strategy Nash Equilibrium

	$Acme: bluray$	$Acme: dvd$
$Best: bluray$	$B = +9;\ A = +9$	$B = -1;\ A = -4$
$Best: dvd$	$B = -1;\ A = -3$	$B = +5;\ A = +5$

- Game: (Classic)
  - Two video game console manufacturers, Acme and Best, need to decide whether to use Blu-ray discs or DVD
  - Profits are positive if they both agree to use the same format, but negative otherwise
  - There is no dominant strategy, so iterative elimination of dominant strategies will fail to find a solution.
- How do we find the solution(s) in this case?

Example: Finding Pure Strategy Nash Equilibrium

	$Acme: bluray$	$Acme: dvd$
$Best: bluray$	$B = +9;\ A = +9$	$B = -1;\ A = -4$
$Best: dvd$	$B = -1;\ A = -3$	$B = +5;\ A = +5$

- There are 2 Nash equilibriums:
  - $(bluray, bluray)$ and (dvd, dvd)
- Choice in this case: $(bluray, bluray)$, a Pareto optimal solution
  - Every game has at least one Pareto-optimal solution, and may have multiple

Finding Pure Strategy Nash Equilibrium

- A simple algorithm to find pure Nash equilibrium:
  - For each column, mark cell if it has the maximum payoff for the row player
    - May be more than one
  - For each row, mark cell if it has the maximum payoff for the column player
    - May be more than one
  - Cells with both row and column marks are pure Nash equilibrium
- There are multiple acceptable solutions, but if agents disagree, both suffer
- There are various ways the agents can coordinate in choosing solutions
  - e.g., by communicating & negotiation
  - Games where agents need to communicate are called coordination games

Mixed Strategies Games

Probabilistic action selection with randomized policy

Example: Game of Chicken

	$Player\ A: continue$	$Player\ A: swerve$
$Player\ B: continue$	$B = -10; A = -10$	$B = 2; A = -2$
$Player\ B: swerve$	$B = -2; A = 2$	$B = 0; A = 0$

- Game: (Terrible!)
  - Two teenagers on opposite ends of a road drive towards each other
  - At the last possible moment, they must decide whether to swerve
  - If one continues while the other swerves, the one who swerves is chicken, while the other is brave
  - If both swerve, then both are OK
  - If both continue – *crash*

Example: Game of Chicken

	$Player\ A: continue$	$Player\ A: swerve$
$Player\ B: continue$	$B = -10; A = -10$	$B = 2; A = -2$
$Player\ B: swerve$	$B = -2; A = 2$	$B = 0; A = 0$

- There are two pure strategy Nash equilibriums
  - $(A:\ continue, B:\ swerve)$
  - $(A:\ swerve, B:\ continue)$
- But there is also a mixed strategy that we will compute

Example: Game of Chicken

[p; 1-p]	$Player\ A: continue$	$Player\ A: swerve$
$Player\ B: continue$	$B = -10; A = -10$	$B = 2; A = -2$
$Player\ B: swerve$	$B = -2; A = 2$	$B = 0; A = 0$

- B plays $continue$:
  - Let B play $continue$ with probability p and swerve with probability $(1 - p)$
  - If A plays $continue$ then the expected payoff for A is:
    $$-10p + 2(1 - p) = -12p + 2$$
  - If A plays $swerve$ then the expected payoff for A is:
    $$-2p + 0(1 - p) = -2p$$
  - It would not matter which action A plays if both the expected payoffs are the same; i.e.,
    $$-12p + 2 = -2p \implies p = 1/5$$

Example: Game of Chicken

	[q; Player A: continue	1-q] Player A: swerve
Player B: continue	$B = -10; A = -10$	$B = 2; A = -2$
Player B: swerve	$B = -2; A = 2$	$B = 0; A = 0$

- A plays $continue$:
  - Let A play $continue$ with probability q and swerve with probability $(1 - q)$
  - If B plays $continue$ then the expected payoff for B is:
$$-10q + 2(1 - q) = -12q + 2$$
  - If B plays $swerve$ then the expected payoff for B is:
$$-2q + 0(1 - q) = -2q$$
  - It would not matter which action B plays if both the expected payoffs are the same; i.e.,
$$-12q + 2 = -2q \Longrightarrow q = 1/5$$

Example: Game of Chicken

	$Player\ A: continue$	$Player\ A: swerve$
$Player\ B: continue$	$B = -10; A = -10$ $\boxed{pq = 1/25}$	$B = 2; A = -2$ $\boxed{p(1-q) = 4/25}$
$Player\ B: swerve$	$B = -2; A = 2$ $\boxed{(1-p)q = 4/25}$	$B = 0; A = 0$ $\boxed{(1-p)(1-q) = 16/25}$

$$p = q = \frac{1}{5}$$

Probabilities of the pairs of actions

Finding mixed strategy:

- For player B the expected utility is:
  - $\frac{1}{25} \times -10 + \frac{4}{25} \times 2 + \frac{4}{25} \times -2 + \frac{16}{25} \times 0 = -\frac{2}{5}$
- For player A the expected utility is:
  - $\frac{1}{25} \times -10 + \frac{4}{25} \times -2 + \frac{4}{25} \times 2 + \frac{16}{25} \times 0 = -\frac{2}{5}$

Expected payoff for each player for this strategy profile

Finding Mixed Strategies

- Algorithm finds a mixture for one player where the other player becomes indifferent to the choices; repeat this for both players
  - Does not always work! (If there is no mixture where the player becomes indifferent)
  - E.g., if one choice dominates the other choice, the player cannot be indifferent to the choices

- Questions:
  - Is there an efficient algorithm for computing a Nash equilibrium in general?

Interpreting Mixed Strategy Equilibria

- What does it mean to play a mixed strategy? (Different interpretations)
  - Randomize to confuse your opponent
    - E.g., Matching pennies
  - Players randomize when they are uncertain about the other's action
    - E.g., Battle of the sexes
  - Mixed strategies are a concise description of what might happen in repeated play
    - Count of pure strategies in the limit
  - Mixed strategies describe population dynamics
    - 2 agents randomly chosen from a population, all having deterministic strategies. Mixed strategy is the probability of getting an agent who will play one pure strategy or another.

Zero-Sum Games

Minimax Theorem

Example: Two-Finger Morra

	$O: one$	$O: two$
$E: one$	$E = +2; O = -2$	$E = -3; O = +3$
$E: two$	$E = -3; O = +3$	$E = +4; O = -4$

- Two players: E and O simultaneously display one or two fingers
- Let, total number of fingers be $f \in \{1, 2, 3, 4\}$:
  - If f is odd, O collects $\$f$ from E
  - If f is even, E collects $\$f$ from O
- What is the optimal mixed strategy against a rational player?
- What is the expected utility for each player?
- Special case of two player zero-sum games: sum of the payoffs is always zero
  - Only need to consider payoff of one player, since the sum is 0: Use the payoff of the even player, E
  - E will try to maximize the payoff $U_E(e, o)$; O will try to minimize the payoff

Efficient algorithm to compute the equilibrium is known

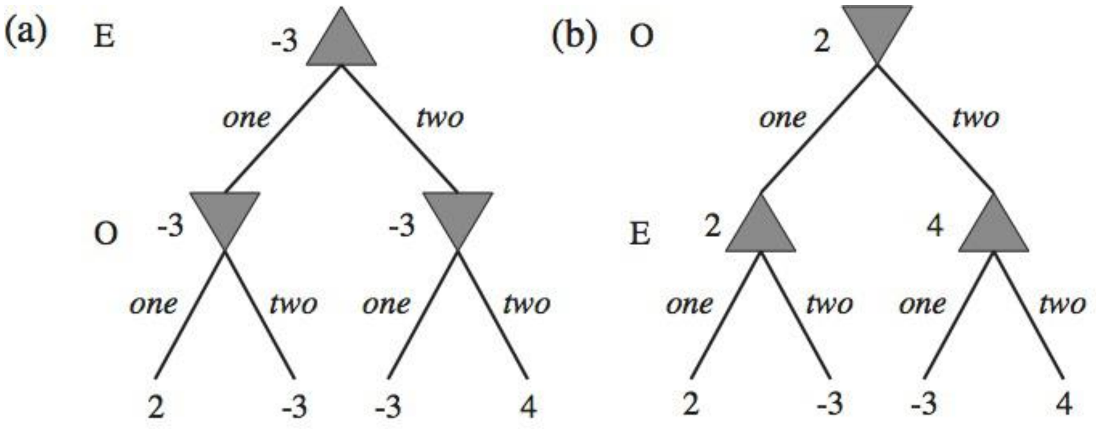
Two Finger Morra: Minimax Game Tree

	$O: one$	$O: two$
$E: one$	$E = +2; O = -2$	$E = -3; O = +3$
$E: two$	$E = -3; O = +3$	$E = +4; O = -4$

- Find range of true utility U of solution
  - Examining pure strategies:

$$U_{E,O} \leq U \leq U_{O,E}: -3 \leq U \leq 2$$

E picks action first & reveals it to O



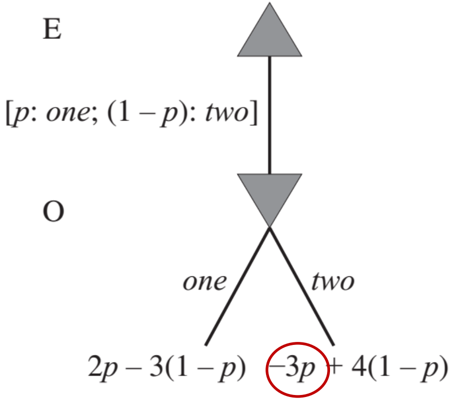
O picks action first & reveals it to E

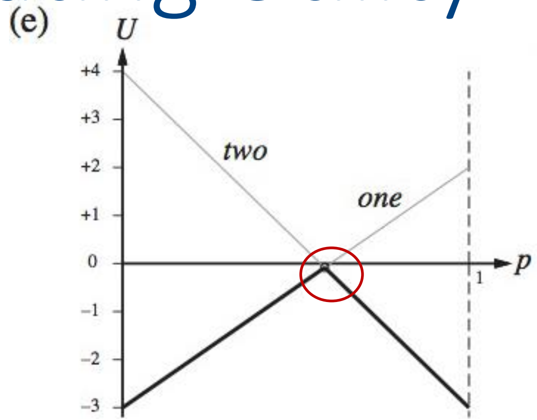
Source: RN Figure 17.12 (a) and (b)

Two Finger Morra: Calculating Utility

- For E:

	one	two
One	$2, -2$	$-3, +3$
Two	$-3, +3$	$+4, -4$





  - If O chooses one: expected payoff to E: $2p - 3(1 - p) = 5p - 3$
  - If O chooses two: expected payoff to E: $-3p + 4(1 - p) = 4 - 7p$
- Draw straight lines in graph (e)
  - O, the minimizer, will always choose lower of the two lines
  - The best E can do at the root is to choose p at intersection pt:

$$5p - 3 = 4 - 7p \implies p = 7/12 \text{ and hence } U_{E,O} = -1/12$$

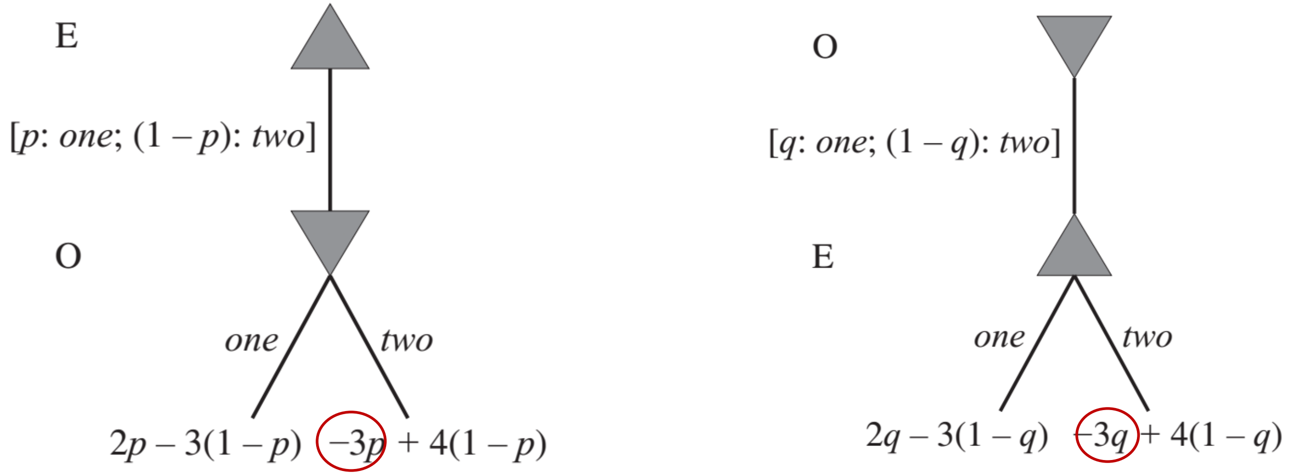
Source: RN Figure 18.12 (c) and (e)

Two Finger Morra: Parameterized Game Tree

	$O: one$	$O: two$
$E: one$	$E = +2; O = -2$	$E = -3; O = +3$
$E: two$	$E = -3; O = +3$	$E = +4; O = -4$

- Determine true value of utility U of solution
  - Examining mixed strategies:

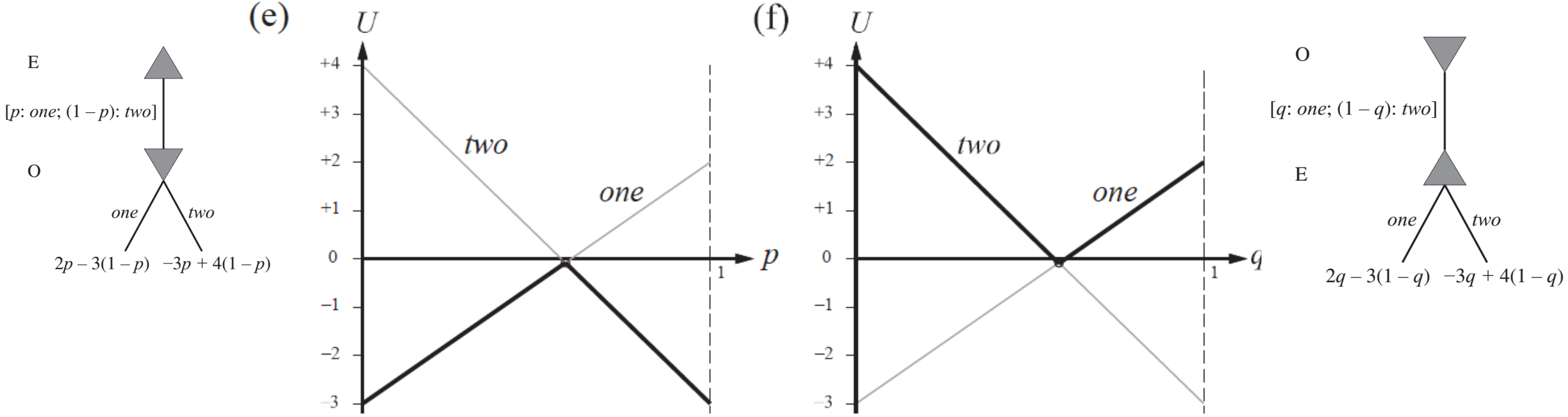
$$U_{E,O} = -1/12, U_{O,E} = -1/12: U = -1/12$$



Source: RN Figure 18.12 (c) and (d)

Two Finger Morra: Utility Profile

- Analyze utility functions to find optimal strategy
  - E always chooses probability parameter for mixed strategy at intersection point

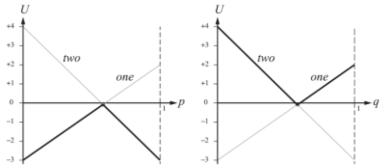


Source: RN Figure 18.12 (e) and (f)

Two Finger Morra: Maximin Equilibrium

- E first: O selects action with smaller expected value, i.e., min of two linear functions.
  - The best that E can do is to select the value p that maximized the min of the two functions (maximin strategy)
  - $p = 7/12$

E first



O first

- O first: E finds max of the two actions
  - The best that O can do is to select the value q that minimizes the function (minimax strategy)
  - $q = 7/12$
- Note: Max of the first function is the same as min of the second! (A special case)
- True utility is attained by the mixed strategy, which should be played by both players.
  - $[7/12: one;\ 5/12: two]$
- This strategy is called the maximin equilibrium of the game, and is a Nash equilibrium.

Minimax Theorem (von Neumann 1928)

- In any finite, two-player, zero-sum game, in any Nash equilibrium each player receives a payoff that is equal to both his maxmin value and his minmax value

- It follows that:
  - Each player's maxmin value is equal to his minmax value. By convention, the maxmin value for player 1 is called the value of the game;
  - For both players, the set of maxmin strategies coincides with the set of minmax strategies; and
  - Any maxmin strategy profile (or minmax strategy profile) is a Nash equilibrium. Furthermore, these are all the Nash equilibria.
    Consequently, all Nash equilibria have the same payoff vector (namely, those in which player 1 gets the value of the game)

Source: [SLB] 3.4.1

Complexity of Computing Nash Equilibrium

- Every finite game is known to have at least one Nash equilibrium
  - For 2-player zero sum game, we know how to find NE efficiently

- However, the problem of finding the NE in the general case is believed to be computationally intractable
  - It is PPAD-complete[1]
    - No known efficient algorithm in general
    - Computationally hard even for 2 player games

[1]https://en.wikipedia.org/wiki/PPAD_(complexity)

Homework

- Readings:
  - RN 17.2.1, 17.2.2
- References:
  - [Ras] E. Rasmussen. Games and Information, 4th ed., Wiley-Blackwell, 2006.
  - http://www.rasmusen.org/GI/download.htm
  - SLB 3.2, 3.3.1-3.3.3, 3.4.1
    [SLB] Y. Shoham & K. Leyton-Brown. Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations. 2009.
    http://www.masfoundations.org/download.html

Repeated Games

(Extra slides)

Repeated Game

- In a repeated game, the players face the same choices repeatedly
  - Each time, history of all players' previous choices is available
- Example: In repeated version of Prisoner's dilemma, Alice & Bob will meet again
  - Q: Will they cooperate instead of testify because of that?

Repeated Game

- In the finite repeated game, we can solve using backward induction
- Consider the version of Prisoner's dilemma where the players know that they will play for exactly 100 rounds
  - On the 100th round, they know that they will play the last game, so will both play the dominant strategy, $testify$
  - Since the strategies for the 100th round is determined, the strategies for 99th round will have no effect on the future, and the dominant strategy for 99th round is also $testify$
  - This reasoning is repeated, and we find that the players will always testify

Perpetual Punishment

	$Alice: testify$	$Alice: refuse$
$Bob: testify$	$B = -5; A = -5$	$B = 0; A = -10$
$Bob: refuse$	$B = -10; A = 0$	$B = -1; A = -1$

- Consider a different repeated version of the Prisoner's dilemma
  - At each round there is a 99% chance that the players will meet again
    - The expected number of rounds is 100
    - Neither player knows when the last round will be
- Cooperative equilibriums are possible under this setting
  - One equilibrium is both players playing the perpetual punishment strategy
    - Both players cooperate and play $refuse$, but if any player plays $testify$, the other player will play $testify$ from then on forever!

Perpetual Punishment

	$Alice: testify$	$Alice: refuse$
$Bob: testify$	$B = -5; A = -5$	$B = 0; A = -10$
$Bob: refuse$	$B = -10; A = 0$	$B = -1; A = -1$

- Expected future payoff of perpetual cooperation is

$$\sum_{t=0}^{\infty} 0.99^t(-1) = -100$$

- Any deviation from NE where both players playing the perpetual punishment strategy

$$0 + \sum_{t=1}^{\infty} 0.99^t(-5) = -495$$

- Not beneficia for either player to ever deviate

Tit-for-tat!

- Another strategy where players start with $refuse$
- Then copy the other player’s previous move for all subsequent moves

- Works well against a wide variety of strategies