Matrix Games

Ronald Parr
CSCI 2951-F
Brown University

With thanks to Vince Conitzer for some content

What is Game Theory? I

- Very general mathematical framework to study situations where multiple agents interact, including:
  - Popular notions of games
  - Everything up to and including multistep, multiagent, simultaneous move, partial information games
  - Example RP & collaborators research: Aiming sensors to catch hiding enemies, assigning guards to posts
  - Can even include negotiating, posturing and uncertainty about the players and game itself
- von Neumann and Morgenstern (1944) was a major launching point for modern game theory
- Nash: Existence of equilibria in general sum games

(wikipedia)

What is game theory? II

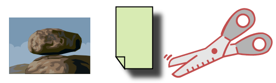
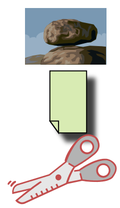
- Study of settings where multiple agents each have
  - Different preferences (utility functions),
  - Different actions
- Each agent's utility (potentially) depends on all agents' actions
  - What is optimal for one agent depends on what other agents do
  - Can be circular
- Game theory studies how agents can rationally form beliefs over what other agents will do, and (hence) how agents should act
- Useful for acting and (potentially) predicting behavior of others
- Not necessarily descriptive

Real World Game Theory Examples

- War
- Auctions
- Animal behavior
- Networking protocols
- Peer to peer networking behavior
- Road traffic
- Related: Mechanism design:
  - Suppose we want people to do X?
  - How to engineer situation so they will act that way?

Rock, Paper, Scissors Zero Sum Formulation

- In zero sum games, one player's loss is other's gain
- Payoff matrix:

	R	P	S
R	0	–1	1
P	1	0	–1
S	–1	1	0

- Minimax solution maximizes worst case outcome

Rock, Paper, Scissors Equations

- R,P,S = probability that we play rock, paper, or scissors respectively (R+P+S = 1)
- U is our expected utility
- Bounding our utility:
  - Opponent rock case: $U \leq P - S$
  - Opponent paper case: $U \leq S - R$
  - Opponent scissors case: $U \leq R - P$
- Want to maximize U subject to constraints
- Solution: (1/3, 1/3, 1/3)

Rock, Paper, Scissors LP Formulation

- Our variables are: $x=[U,R,P,S]^T$
- We want:
  - Maximize U
  - U ≤ P – S
  - U ≤ S – R
  - U ≤ R – P
  - R+P+S = 1
- How do we make this fit:

$$\begin{aligned} \text{maximize}: & \ c^T x \\ \text{subject to}: & \ \mathbf{A}x \le b \\ : & \ x \ge 0 \end{aligned}$$

?

Rock Paper Scissors LP Formulation

$$x = [U,R,P,S]^T$$

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 1 & 1 & 0 & -1 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{pmatrix}$$

$$b = [0,0,0,1,-1]^T$$

$$c = [1,0,0,0]^T$$

$$\begin{aligned} \text{maximize}: & \ c^T x \\ \text{subject to}: & \ \mathbf{A}x \le b \\ : & \ x \ge 0 \end{aligned}$$

First row of Ax: U – P + S ≤ 0

Rock, Paper, Scissors Solution

- If we feed this LP to an LP solver we get:
  - R=P=S=1/3
  - U=0
- Solution for the other player is:
  - The same…
  - By symmetry
- This is the minimax solution
- This is also an equilibrium
  - No player has an incentive to deviate
  - (Defined more precisely later)

Tangent: Why is RPS Fun?

- OK, it's not…

- Why *might* RPS be fun?
  - Try to exploit non-randomness in your friends
  - Try to be random yourself

Generalizing

- We can solve any two player, simultaneous move, zero sum game with an LP
  - One variable for each of player 1's actions
  - Variables must be a probability distribution (constraints)
  - One constraint for each of player 2's actions (Player 1's utility must be less than or equal to outcome for each player 2 action.)
  - Maximize player 1's utility
- Can solve resulting LP using an LP solver in time that is (weakly) polynomial in total number of actions

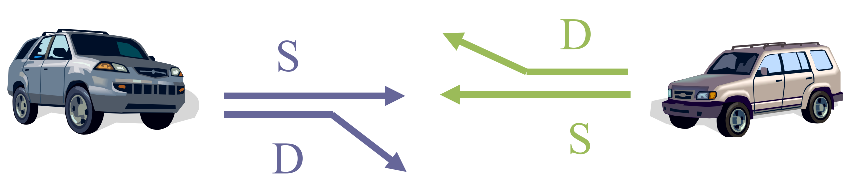
Minimax Solutions in General

- What do we know about minimax solutions?
  - Can a suboptimal opponent trick minimax?
  - When should we abandon minimax?
- Minimax solutions for 2-player zero-sum games can always be found by solving a linear program
- The minimax solutions will also be equilibria (more on that later)
- For general sum games:
  - Minimax does not apply
  - Solutions (equilibria) may not be unique
  - Search for equilibria using more computationally intensive methods

General Sum Games

"Chicken"

- Two players drive cars towards each other
- If one player goes straight, that player wins
- If both go straight, they both die



Source: wikipedia

	D	S
D	0, 0	-1, 1
S	1, -1	-5, -5

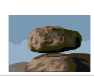
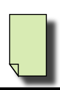
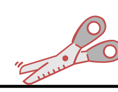
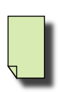
not zero-sum

Reasoning About General Sum Games

- Can’t approach as an optimization problem
- Minimax doesn’t apply
  - Other players’ objectives might be aligned w/ yours
  - Might be partially aligned
- Need a solution concept where each players is “satisfied” WRT his/her objectives

Rock-paper-scissors – Seinfeld variant

MICKEY: All right, rock beats paper!
(Mickey smacks Kramer's hand for losing)
KRAMER: I thought paper covered rock.
MICKEY: Nah, rock flies right through paper.
KRAMER: What beats rock?
MICKEY: (looks at hand) Nothing beats rock.

			
	0, 0	1, -1	1, -1
	-1, 1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

Note: still zero-sum, but useful for understanding a different way of thinking about game solutions.

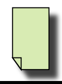
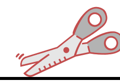
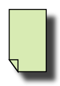
Dominance

- Player i's strategy s_i strictly dominates s_i' if
  - for any s_{-i}, $u_i(s_i, s_{-i}) > u_i(s_i', s_{-i})$
- s_i weakly dominates s_i' if
  - for any s_{-i}, $u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$; and
  - for some s_{-i}, $u_i(s_i, s_{-i}) > u_i(s_i', s_{-i})$

-i = "the player(s) other than i"

strict dominance

weak dominance

			
	0, 0	1, -1	1, -1
	-1, 1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

Prisoner's Dilemma

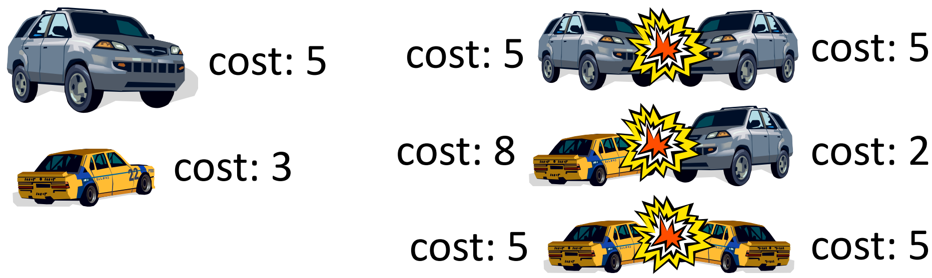
- Pair of criminals has been caught
- District attorney has evidence to convict them of a minor crime (1 year in jail); knows that they committed a major crime together (3 years in jail) but cannot prove it
- Offers them a deal:
  - If both confess to the major crime, they each get a 1 year reduction
  - If only one confesses, that one gets 3 years reduction

	confess	don't confess
confess	-2, -2	0, -3
don't confess	-3, 0	-1, -1

“Should I buy an SUV?”

purchasing + gas cost

accident cost

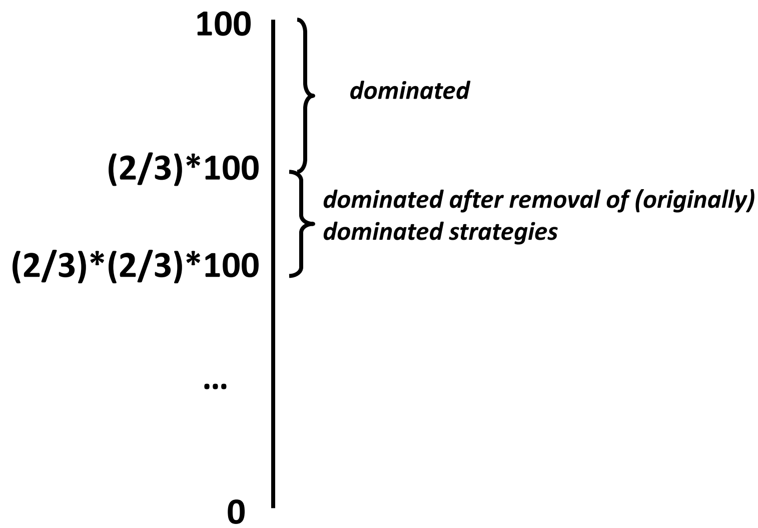


		
	-10, -10	-7, -11
	-11, -7	-8, -8

“2/3 of the average” game

- Everyone writes down a number between 0 and 100
- Person closest to 2/3 of the average wins
- Example:
  - A says 50
  - B says 10
  - C says 90
  - Average(50, 10, 90) = 50
  - 2/3 of average = 33.33
  - A is closest (|50-33.33| = 16.67), so A wins

"2/3 of the average" game revisited



Iterated dominance

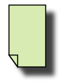
- Iterated dominance: remove (strictly/weakly) dominated strategy, repeat
- Iterated strict dominance on Seinfeld's RPS:

0, 0	1, -1	1, -1
-1, 1	0, 0	-1, 1
-1, 1	1, -1	0, 0

→

0, 0	1, -1
-1, 1	0, 0

Mixed strategies

- Mixed strategy for player i = probability distribution over player i's (pure) strategies
- E.g. 1/3  1/3 , 1/3 
- Example of dominance by a mixed strategy:

1/2	3, 0	0, 0
1/2	0, 0	3, 0
	1, 0	1, 0

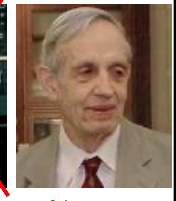
Best Responses

- Let A be a matrix of player 1's payoffs
- Let σ_2 be a mixed strategy for player 2
- $A\sigma_2$ = vector of expected payoffs for each strategy for player 1
- Highest entry indicates **best response** for player 1
- Any mixture of ties is also BR, but *can only tie a pure BR*
- Generalizes to >2 players

0, 0	-1, 1
1, -1	-5, -5

σ_2

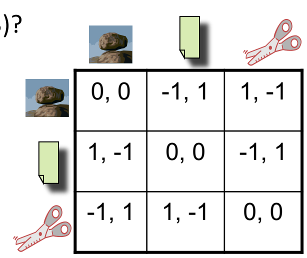
Nash equilibrium [Nash 50]



- A vector of strategies (one for each player) = a strategy profile
- Strategy profile ($\sigma_1, \sigma_2, \ldots, \sigma_n$) is a Nash equilibrium if each σ_i is a best response to σ_{-i}
  - That is, for any i, for any σ_i', $u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$
- Does not say anything about multiple agents changing their strategies at the same time
- In any (finite) game, at least one Nash equilibrium (possibly using mixed strategies) exists [Nash 50]
- (Note - singular: equilibrium, plural: equilibria)

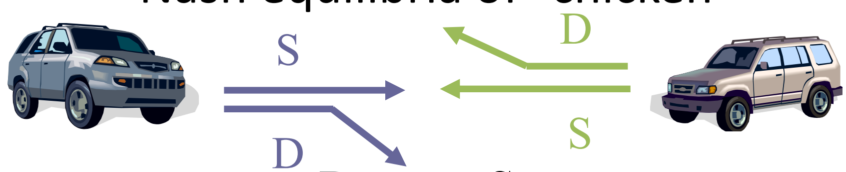
Equilibrium Strategies vs. Best Responses

- equilibrium strategy -> best response?
- best response -> equilibrium strategy?
- Consider Rock-Paper-Scissors
  - Is (1/3, 1/3, 1/3) a best response to (1/3, 1/3, 1/3)?
  - Is (1, 0, 0) a best response to (1/3, 1/3, 1/3)?
  - Is (1, 0, 0) a strategy for any equilibrium?



0, 0	-1, 1	1, -1
1, -1	0, 0	-1, 1
-1, 1	1, -1	0, 0

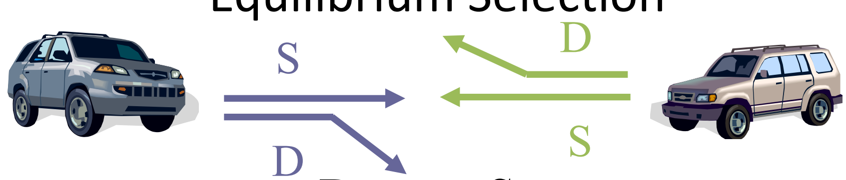
Nash equilibria of "chicken"



	D	S
D	0, 0	-1, 1
S	1, -1	-5, -5

- (D, S) and (S, D) are Nash equilibria
  - They are pure-strategy Nash equilibria: nobody randomizes
  - They are also strict Nash equilibria: changing your strategy will make you strictly worse off
- No other pure-strategy Nash equilibria

Equilibrium Selection



	D	S
D	0, 0	-1, 1
S	1, -1	-5, -5

- (D, S) and (S, D) are Nash equilibria
- Which do you play?
- What if player 1 assumes (S, D), player 2 assumes (D, S)
- Play is (S, S) = (-5, -5)!!!

- This is the **equilibrium selection** problem

Nash equilibria of "chicken"…

	D	S
D	0, 0	-1, 1
S	1, -1	-5, -5

- Is there a Nash equilibrium that uses mixed strategies -- say, where player 1 uses a mixed strategy?
- ***If a mixed strategy is a best response, then all of the pure strategies that it randomizes over must also be best responses***
- So we need to make player 1 indifferent between D and S
- Player 1's utility for playing D = $-p^c_S$
- Player 1's utility for playing S = $p^c_D - 5p^c_S = 1 - 6p^c_S$
- So we need $-p^c_S = 1 - 6p^c_S$ which means $p^c_S = 1/5$
- Then, player 2 needs to be indifferent as well
- Mixed-strategy Nash equilibrium: ((4/5 D, 1/5 S), (4/5 D, 1/5 S))
  - People may die! Expected utility -1/5 for each player

$-p^c_S$ = probability that column player plays s

Does This Technique Generalize?

- Sort of…
- For two players:
  - If you guess which actions have non-zero probability in equilibrium
  - Can solve for equilibrium probabilities
  - (exponential time in worst case)
- For >2 players, things get more complicated
- Searching through all subsets of actions is exponential, but
  - Iterating in order of increasing support works surprisingly well
  - Why? Empirically, many games with large action spaces often have equilibria using only a small(ish) number of actions

Strategy Generation (Double Oracle)

- Assumptions:
  - You can afford to solve small games (small # of actions)
  - You can efficiently compute a best response with a best response oracle (not a crazy assumption)
- Double Oracle Algorithm:
  - Initialize each player with one available action each
  - Repeat
    - Compute equilibrium
    - Compute best responses
    - If best responses use actions already available, return equilibrium
    - Else, add best responses to set of available actions
- Guaranteed to converge, often w/o using all possible actions
- For zero-sum games, Double Oracle can be viewed as an instance of constraint generation for linear programs

NE as a Non-linear Program

$$U^i(\boldsymbol{\pi}) = \sum_{\mathbf{a} \in \mathcal{A}} R^i(\mathbf{a}) \prod_{j \in \mathcal{I}} \pi^j(a^j)$$

← Utility of joint policy π from perspective of player i

$$\begin{aligned} \underset{\boldsymbol{\pi}, U}{\text{minimize}} \quad & \sum_i \left(U^i - U^i(\boldsymbol{\pi}) \right) \\ \text{subject to} \quad & U^i \geq U^i(a^i, \boldsymbol{\pi}^{-i}) \text{ for all } i, a^i \\ & \sum_{a^i} \pi^i(a^i) = 1 \text{ for all } i \\ & \pi^i(a^i) \geq 0 \text{ for all } i, a^i \end{aligned}$$

U^i terms are nonlinear ☹

Other formulations exist, but none are polynomial time

Computational Issues

- Zero-sum games - solved efficiently as LP
- Equilibria of general sum games are guaranteed to exist (Nash), but may require exponential time (in # of actions) to find a single equilibrium
- Determining whether an equilibrium exists that has certain properties (e.g., utility > x for player i?) is NP-hard
- Producing any equilibrium is PPAD complete
  (PPAD is like NP, but problems are not decision problems.)
- Despite bad worst-case, many games solved with existing algorithms
- Many tractable special cases exist

Other Approaches (not guaranteed to converge)

- Iterated best response
  - Iterate over players in some (random) order
  - Adopt best response strategy given other players' strategies
  - Converges in some cases
  - Limiting behavior (avg. of best responses) may approximate NE
- Fictitious play
  - Estimate opponent stochastic strategies averaging over previous plays
  - Play best response to opponent strategies
  - Repeat for all players
  - Converges in some cases
- These methods can get into cycles

Correlated Equilibrium

- So far, assumed agents choose actions independently, i.e., probability of joint action is product of probabilities of individual actions
- Correlated equilibrium (CE) allows joint action distribution to be an arbitrary distribution
  - Pros:
    - More natural model in some problems
    - Compute in poly time in number of parameters of the distribution
  - Cons:
    - Correlation mechanism not natural for some domains
    - Size of joint distribution over action space of all agents can be large
      - NE for n agents with k actions is nxk numbers
      - CE for n agents with k actions is k^n numbers
        (though same can be said about payoff matrix itself)

Computing CE

a is a vector, so this is a sum over all possible joint actions

$$\underset{\pi}{\text{maximize}} \quad \sum_i \sum_{\mathbf{a}} R^i(\mathbf{a})\pi(\mathbf{a})$$

$$\text{subject to} \quad \sum_{\mathbf{a}^{-i}} R^i(a^i, \mathbf{a}^{-i})\pi(a^i, \mathbf{a}^{-i}) \geq \sum_{\mathbf{a}^{-i}} R^i(a^{i\prime}, \mathbf{a}^{-i})\pi(a^i, \mathbf{a}^{-i}) \quad \text{for all } i, a^i, a^{i\prime}$$

$$\sum_{\mathbf{a}} \pi(\mathbf{a}) = 1$$

$$\pi(\mathbf{a}) \geq 0 \quad \text{for all } \mathbf{a}$$

- Good news: This is a linear program
- Bad news: Can have a large number of variables
- NE ⊆ CE

Game Theory Issues

- How descriptive is game theory?
  - Some evidence that people play equilibria
  - Also, some evidence that people act irrationally
  - If it is computationally intractable to solve for equilibria of large games, seems unlikely that people are doing this
- How reasonable is (basic) game theory?
  - Are payoffs known?
  - Are situations really simultaneous-move with no information about how the other player will act?
  - Are situations really single-shot? (repeated games?)
  - How is equilibrium selection handled in practice?

Extensions

- Partial information
- Uncertainty about the game parameters, e.g., payoffs (Bayesian games)
- Repeated games: Simple learning algorithms can converge to equilibria in some repeated games
- Multistep games with distributions over next states (game theory + MDPs = stochastic games)
- Multistep + partial information (Partially observable stochastic games)

- Game theory is so general, that it can encompass essentially all aspects of strategic, multiagent behavior, e.g., negotiating, threats, bluffs, coalitions, bribes, etc.

Conclusions

- Game theory tells us how to act in strategic situations – different agents with different goals acting with awareness of other agents
- Zero sum case is relatively easy
- General sum case is computationally hard – though some nice results exist for special cases
- Extensions address some shortcomings/assumptions of basic model but at additional computational cost