A Tutorial on Computational Game Theory

NIPS 2002

Michael Kearns

Computer and Information Science

University of Pennsylvania

mkearns@cis.upenn.edu

For an updated and expanded version of these slides, visit http://www.cis.upenn.edu/∼mkearns/nips02tutorial

Thanks To:

- Avrim Blum
- Dean Foster
- Sham Kakade
- Jon Kleinberg
- Daphne Koller
- John Langford
- Michael Littman
- Yishay Mansour
- Andrew Ng
- Luis Ortiz
- David Parkes
- Lawrence Saul
- Rob Schapire
- Yoav Shoham
- Satinder Singh
- Moshe Tennenholtz
- Manfred Warmuth

Road Map (1)

- Examples of Strategic Conflict as Matrix Games
- Basics Definitions of (Matrix) Game Theory
- Notions of Equilibrium: Overview
- Definition and Existence of Nash Equilibria
- Computing Nash Equilibria for Matrix Games
- Graphical Models for Multiplayer Game Theory
- Computing Nash Equilibria in Graphical Games

Road Map (2)

- Other Equilibrium Concepts:
  - Correlated Equilibria
  - Correlated Equilibria and Graphical Games
  - Evolutionary Stable Strategies
  - Nash's Bargaining Problem, Cooperative Equilibria
- Learning in Repeated Games
  - Classical Approaches; Regret Minimizing Algorithms
- Games with State
  - Connections to Reinforcement Learning
- Other Directions and Conclusions

Example: Prisoner's Dilemma

- Two suspects in a crime are interrogated in separate rooms
- Each has two choices: confess or deny
- With no confessions, enough evidence to convict on lesser charge; one confession enough to establish guilt
- Police offer plea bargains for confessing
- Encode strategic conflict as a payoff matrix:

payoffs	confess	deny
confess	$-3,-3$	$0,-4$
deny	$-4,0$	$-1,-1$

- What should happen?

Example: Hawks and Doves

- Two players compete for a valuable resource
- Each has a confrontational strategy ("hawk") and a conciliatory strategy ("dove")
- Value of resource is V; cost of losing a confrontation is C
- Suppose $C > V$ (think nuclear first strike)
- Encode strategic conflict as a payoff matrix:

payoffs	hawk	dove
hawk	$(V-C)/2$,$(V-C)/2$	V,0
dove	0,V	$V/2$,$V/2$

- What should happen?

A (Weak) Metaphor

- Actions of the players can be viewed as (binary) variables
- Under any reasonable notion of “rationality”, the payoff matrix imposes constraints on the joint behavior of these two variables
- Instead of being probabilistic, these constraints are strategic
- Instead of computing conditional distributions given the other actions, players optimize their payoff
- Players are selfish and play their best response

Basics of Game Theory

- Set of players $i = 1, \ldots, n$ (assume $n = 2$ for now)
- Each player has a set of m basic actions or pure strategies (such as “hawk” or ”dove”)
- Notation: a_i will denote the pure strategy chosen by player i
- Joint action: $\vec{a}$
- Payoff to player i given by matrix or table $M_i(\vec{a})$
- Goal of players: maximize their own payoff

Notions of Equilibria: Overview (1)

- An equilibrium among the players is a strategic standoff
- No player can improve on their current strategy
- But under what model of communication, coordination, and collusion among the players?
- All standard equilibrium notions are descriptive rather than prescriptive

Notions of Equilibria: Overview (2)

- No communication or bargaining:
  Nash Equilibria
- Communication via correlation or shared randomness:
  Correlated Equilibria
- Full communication and coalitions:
  (Assorted) Cooperative Equilibria
- Equilibrium under evolutionary dynamics:
  Evolutionary Stable Strategy
- We'll begin with Nash Equilibria

Mixed Strategies

- Need to introduce mixed strategies
- Each player i has an independent distribution p_i over their pure strategies ($p_i \in [0, 1]$ in 2-action case)
- Use $\vec{p} = (p_1, \ldots, p_n)$ to denote the product distribution induced over joint action $\vec{a}$
- Use $\vec{a} \sim \vec{p}$ to indicate $\vec{a}$ distributed according to $\vec{p}$
- Expected return to player i: $\mathbf{E}_{\vec{a} \sim \vec{p}}[M_i(\vec{a})]$
- (What about more general distributions over $\vec{a}$?)

Nash Equilibria

- A product distribution $\vec{p}$ such that no player has a unilateral incentive to deviate
- All players know all payoff matrices
- Informal: no communication, deals or collusion allowed — everyone for themselves
- Let $\vec{p}[i : p'_i]$ denote $\vec{p}$ with p_i replaced by p'_i
- Formally: $\vec{p}$ is a Nash equilibrium (NE) if for every player i, and every mixed strategy p'_i, $\mathbf{E}_{\vec{a}\sim\vec{p}}[M_i(\vec{a})] \geq \mathbf{E}_{\vec{a}\sim\vec{p}[i:p'_i]}[M_i(\vec{a})]$
- Nash 1951: NE always exist in mixed strategies
- Players can announce their strategies

Approximate Nash Equilibria

- A set of mixed strategies $(\vec{p}_1, \ldots, \vec{p}_n)$ such that no player has "too much" unilateral incentive to deviate
- Formally: $\vec{p}$ is an ϵ-Nash equilibrium (NE) if for every player i, and every mixed strategy p'_i, $\mathbf{E}_{\vec{a}\sim\vec{p}}[M_i(\vec{a})] \geq \mathbf{E}_{\vec{a}\sim\vec{p}[i:p'_i]}[M_i(\vec{a})] - \epsilon$
- Motivation: intertia, cost of change,. . .
- Computational advantages

NE for Prisoner's Dilemma

- Recall payoff matrix:

payoffs	confess	deny
confess	$-3,-3$	$0,-4$
deny	$-4,0$	$-1,-1$

- One (pure) NE: (confess,confess)
- Failure to cooperate despite benefits
- Source of great and enduring angst in game theory

NE for Hawks and Doves

- Recall payoff matrix ($V < C$):

payoffs	hawk	dove
hawk	$(V - C)/2$,$(V - C)/2$	V,0
dove	0,V	$V/2$,$V/2$

- Three NE:
  - pure: (hawk,dove)
  - pure: (dove,hawk)
  - mixed: (Pr[hawk] $= V/C$,Pr[hawk] $= V/C$)
- Rock-Paper-Scissors: Only mixed NE

NE Existence Intuition

- Suppose that $\vec{p}$ is not a NE
- For some player i, must be some pure strategy giving higher return against $\vec{p}$ than p_i
- For each such player, shift some of the weight of p_i to this pure strategy
- Leave all other p_j alone
- Formalize as continuous mapping $\vec{p} \rightarrow F(\vec{p})$
- Brouwer Fixed Point Theorem: continuous mapping F of a compact set into itself must possess $\vec{p}^*$ such that $F(\vec{p}^*) = \vec{p}^*$
- One-dimensional case easy, high-dimensional difficult

Some NE Facts

- Existence not guaranteed in pure strategies
- May be multiple NE
- In multiplayer case, may be exponentially many NE
- Suppose (p_1, p_2) and (p'_1, p'_2) are two NE
- Zero-sum: (p_1, p'_2) and (p'_1, p_2) also NE, and give players same payoffs (games have a unique value)
- General sum: (p_1, p'_2) may not be a NE; different NE may give different payoffs
- Which will be chosen?
  - dynamics, additional criteria, structure of interaction?

Computing NE

- Inputs:
  - Payoff matrices M_i
  - Note: each has m^n entries (n players, m actions each)
- Output:
  - Any NE?
  - All NE? (output size)
  - Some particular NE?

Complexity Status of Computing a NE (1)

- Zero-sum, 2-player case (input size m^2):
  - Linear Programming
  - Polynomial time solution
- General-sum case, 2 players (input size m^2):
  - Closely related to Linear Complementarity Problems
  - Can be solved with the Lemke-Howson algorithm
  - Exponential worst-case running time
  - Probably not in P, but probably not NP-complete?

Complexity Status of Computing a NE (2)

- Maximizing sum of rewards NP-complete for 2 players
- General-sum case, multiplayer (input size m^n):
  - Simplical subdivision methods (Scarf's algorithm)
  - Exponential worst-case running time
  - Not clear small action spaces ($n = 2$) help
- Missing: compact models of large player and action spaces

2-Player, Zero-Sum Case: LP Formulation

- Assume 2 players, $M = M_1 = -M_2$
- Let $p_1 = (p_1^1, \ldots, p_1^m)$ and p_2 be mixed strategies
- Minimax theorem says:

$$\max_{p_1} \min_{p_2} \{p_1 M p_2\} = \min_{p_2} \max_{p_1} \{p_1 M p_2\}$$

- Solved by standard LP methods

General Sum Case: A Sampling Folk Theorem

- Suppose (p_1, p_2) is a NE
- Idea: let $\widehat{p}_i$ be an empirical distribution by sampling p_i
- If we sample enough, $\widehat{p}_i$ and p_i will get nearly identical returns against any opponent strategy (uniform convergence)
- Thus, $(\widehat{p}_1, \widehat{p}_2)$ will be ϵ-NE
- From Chernoff bounds, only $\approx (1/\epsilon^2)\log(m)$ samples suffices
- Yields $(m)^{(1/\epsilon^2)\log(m)}$ algorithm for approximate NE

Compact Models for Multiplayer Games

- Even in 2-player games, computational barriers appear
- Multiplayer games make things even worse
- Maybe we need better representations
- *See accompanying PowerPoint presentation.*

Correlated Equilibria

- NE $\vec{p}$ is a product distribution over the joint action $\vec{a}$
- Suffices to guarantee existence of NE
- Now let P be an arbitrary joint distribution over $\vec{a}$
- Informal intuition: assuming all others play "their part" of P, i has no unilateral incentive to deviate from P
- Let $\vec{a}_{-i}$ denote all actions except a_i
- Say that P is a Correlated Equilibrium (CE) if for any player i, and any actions a, a' for i:

$$\sum_{\vec{a}_{-i}} P(\vec{a}_{-i}|a_i = a) M_i(\vec{a}_{-i}, a) \geq \sum_{\vec{a}_{-i}} P(\vec{a}_{-i}|a_i = a) M_i(\vec{a}_{-i}, a')$$

Advantages of CE

- Conceptual: Some CE payoff vectors not achievable by NE
- Everyday example: traffic signal
- CE allows "cooperation" via shared randomization
- Any mixture of NE is a CE — but there are other CE as well
- Computational: note that

  $$\sum_{\vec{a}_i}(P(\vec{a}_{-i}, a_i = a)/P(a_i = a))M_i(\vec{a}_{-i}, a) \geq \sum_{\vec{a}_i}(P(\vec{a}_{-i}, a_i = a)/P(a_i = a'))M_i(\vec{a}_{-i}, a)$$

  is linear in variables $P(\vec{a}_{-i}, a_i = a) = P(\vec{a})$
- Thus have just a linear feasibility problem
- 2-player case: compute CE in polynomial time

Correlated Equilibria and Graphical Games

- No matter how complex the game, NE factor
- Thus, NE always have compact representations
- Any mixture of NE is a CE
- Thus, even simple games can have CE of arbitrary complexity
- How do we represent the CE of a graphical game?
- Restrict attention to CE up to expected payoff equivalence

Markov Nets and Graphical Games

- Let G be the graph of a graphical game
- Can define a Markov net $MN(G)$:
  - Form cliques of local neighborhoods in G
  - For each clique C, introduce potential function $\phi_C \geq 0$ on just the settings in C
  - Markov net semantics: $\Pr[\vec{a}] = (1/Z)\prod_C \phi(\vec{a}_C)$
- For any CE of a game with graph G, there is a CE with identical expected payoffs representable in $MN(G)$
- Link between strategic and probabilistic structure
- If G is a tree, can compute a (random) CE efficiently

Evolutionary Game Theory

- A different model of multiplayer games
- Assume an infinite population of players — but that meet in random, pairwise confrontations
- Assume symmetric payoff matrix M (as in Hawks and Doves)
- Let P be the distribution over actions induced by the (averaged) population mixed strategies p_i
- Then fitness of p_i is expected return against P
- Assume evolutionary dynamics: the higher the fitness of p_i, the more offspring player i has in the next generation

Evolutionary Stable Strategies

- Let P be the population mixed strategy
- Let Q be an invading "mutant" population
- Let $M(P,Q)$ be the expected payoff to a random player from P facing a random player from Q
- Suppose population is $(1-\epsilon)P+\epsilon Q$
- Fitness of incumbent population: $(1-\epsilon)M(P,P)+\epsilon M(P,Q)$
- Fitness of invading population: $(1-\epsilon)M(Q,P)+\epsilon M(Q,Q)$
- Say P is an ESS if for any $Q \neq P$ and sufficiently small $\epsilon > 0$,
$$(1-\epsilon)M(P,P)+\epsilon M(P,Q) > (1-\epsilon)M(Q,P)+\epsilon M(Q,Q)$$
- Either $M(P,P) > M(Q,P)$ or $M(P,P) = M(Q,P)$ and $M(P,Q) > M(Q,Q)$

ESS for Hawks and Doves

- Recall payoff matrix ($V < C$):

payoffs	hawk	dove
hawk	$(V-C)/2$,$(V-C)/2$	V,0
dove	0,V	$V/2$,$V/2$

- ESS: $P(\text{hawk}) = V/C$

Remarks on ESS

- Do not always exist!
- Special type of (symmetric) NE
- Biological field studies
- Sources of randomization
- Mixed strategies vs. population averages
- Market models

Richer Game Representations

- Have said quite a lot about single-shot matrix games
- What about:
  - Repeated games
  - Games with state (chess, checkers)
  - Stochastic games (multi-player MDPs)
- Can always (painfully) express in normal form
- Normal form equilibria concepts relevant

Repeated Games

- Still have underlying game matrices
- Now play the single-shot game repeatedly, examine cumulative or average reward
- Game has no internal state (though players might)
- Relevant detail: how many rounds of play?

Learning in Repeated Games

- "Classical" algorithms:
  - Fictitious Play: best response to empirical distribution of opponent play
  - Various (stochastic) gradient approaches
- Common question: when will such dynamics converge to NE?
- Positive results fairly restrictive
- Generalizations to parametric strategy representations?

Exponential Updates and Regret Minimization

- View repeated play as a sequence of trials against an arbitrary opponent
- Maintain a weight on each pure strategy
- On each trial, multiply each weight by a factor exponentially decreasing in its regret
- General setting: near-minimization of regret on sequence, but no guarantee of NE
- Zero-sum case: two “copies” will converge to NE
- Regret minimization and NE vs. CE

Repeated Games and Bounded Rationality

- Consider restricting the complexity of strategies in T rounds of a repeated game
- Example: next action computed by a finite state machine on the history of play so far
- New equilibria may arise from the restriction
- Prisoner's Dilemma: if number of states is $o(\log(T))$, mutual cooperation (denial) becomes a NE

Games with State

- Standard board games: chess, checkers
- Often feature partial or hidden information (poker)
- Might involve randomization (backgammon)

Stochastic Games

- Generalize MDPs to multiple players
- At each state s, have payoff matrix M_i^s for player i
- Immediate reward to i at state s under joint action $\vec{a}$ is $M_i^s(\vec{a})$
- Markovian dynamics: $P(s'|s,\vec{a})$
- Discounted sum of rewards
- Every player has a policy $\pi_i(s)$
- Generalize optimal policy to (Nash) equilibrium $(\pi_1,\ldots,\pi_n)$
- Don't just have to worry about influence on future state, but everyone else's policy
- Exploration even more challenging

Stochastic Games and RL

- For fixed policies of opponents, can define value functions
- What happens when independent Q-learners play?
- Results with different amounts and type of shared info
- Generalization of E^3 algorithm to stochastic games
- Generalization of sparse sampling methods

Conclusions

- Classical game theory a rich and varied formalism for strategic reasoning, a complement to more passive reasoning
- Like probability theory, provides sound foundations but lacks emphasis on representation and computation
- Computational game theory aims to provide these emphases
- Many substantive connections to NIPS topics already under way (graphical models, learning algorithms, dynamical systems, reinforcement learning)...
- ... but even more lie ahead.
- Come find me to chat about open problems!

Contact Information

- Email: mkearns@cis.upenn.edu
- Web: www.cis.upenn.edu/~mkearns
- This tutorial: www.cis.upenn.edu/~mkearns/nips02tutorial
  - will morph into Penn course page
- COLT/SVM 2003 special session on game theory