A Glimpse of Game Theory 08
Games and Game Theory

• Much effort to develop computer programs for artificial games like chess or poker commonly played for entertainment

• Larger issue: account for, model, and predict how agents (human or artificial) interact with other agents

• **Game theory** accounts for mixture of cooperative and competitive behavior

• Applies to zero-sum and non-zero-sum games
Basic Ideas of Game Theory

- **Game theory** studies how strategic interactions among **rational players** produce **outcomes** with respect to players’ **preferences**
  - Preferences represented as utilities (numbers)
  - Outcomes might not have been intended
- Provides a general theory of strategic behavior
- Generally depicted in mathematical form
- Plays important role in economics, decision theory and **multi-agent systems**
Zero Sum Games

• Zero-sum: participant’s gain/loss exactly balanced by losses/gains of the other participants

• Total gains of participants minus total losses = 0
  Poker is zero sum game: money won = money lost

• Commercial trade not a zero-sum game
  If country with an excess of bananas trades with another for their excess of apples, both may benefit

• Non-zero-sum games more complex to analyze
• More non-zero-sum games as world becomes more complex, specialized, and interdependent
Rules, Strategies, Payoffs & Equilibrium

Situations are treated as “games”:

• **Rules** of game: who can do what, and when they can do it

• Player's **strategy**: plan for actions in each possible situation in the game

• Player's **payoff**: amount that player wins or loses in particular situation in a game

• Player has a **dominant strategy** if her best strategy doesn’t depend on what others do
Game Theory Roots

• Defined by John von Neumann & Oskar Morgenstern

• Provides powerful model & practical tools to model interactions among sets of autonomous agents

• Used to model strategic policies (e.g., arms race among countries)
Nash Equilibrium

• Occurs when each player's strategy is optimal given strategies of other players
• It means that no player benefits by unilaterally changing strategy, while others stay fixed
• Like a local maximum in hill climbing
• Every finite game has at least one Nash equilibrium in either pure or mixed strategies (proved by John Nash)
  – Nash won 1994 Nobel Prize in economics for this work
  – Read *A Beautiful Mind* by Sylvia Nasar (1998) and/or see the 2001 film
Prisoner's Dilemma

• Famous example from game theory

• Strategies must be undertaken without full knowledge of what other players will do

• Players adopt dominant strategies, but they don't necessarily lead to the best outcome

• Rational behavior leads to a situation where everyone is worse off!
Bonnie and Clyde

Bonnie and Clyde are arrested and charged with crimes. They’re questioned separately, unable to communicate. They know how it works:

– If both proclaim mutual innocence (cooperating), they will be found guilty anyway and get three-year sentences for robbery
– If one confesses (defecting) and the other doesn’t (cooperating), the confessor is rewarded with a light, one-year sentence and the other gets a severe eight-year sentence
– If both confess (defecting), then the judge sentences both to a moderate four years in prison

What should Bonnie do? What should Clyde do?
The payoff matrix

<table>
<thead>
<tr>
<th>BONNIE</th>
<th>CLYDE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Confess</strong></td>
<td>Confess</td>
</tr>
<tr>
<td><strong>4 years each</strong></td>
<td>1 year for Bonnie</td>
</tr>
<tr>
<td></td>
<td>and 8 years for</td>
</tr>
<tr>
<td></td>
<td>Clyde</td>
</tr>
<tr>
<td><strong>Not Confess</strong></td>
<td>8 years for Bonnie</td>
</tr>
<tr>
<td></td>
<td>and 1 year for Clyde</td>
</tr>
<tr>
<td></td>
<td>3 years each</td>
</tr>
</tbody>
</table>
Bonnie’s Decision Tree

There are two cases to consider

If Clyde Confesses

- **Bonnie**
  - **Confess**: 4 Years in Prison
  - **Not Confess**: 8 Years in Prison

If Clyde Does Not Confess

- **Bonnie**
  - **Confess**: 1 Year in Prison
  - **Not Confess**: 3 Years in Prison

Bonnie’s **Dominant strategy** is to confess (defect) because no matter what Clyde does, she is better off confessing.
So what?

- Clyde’s reasoning is the same
  - They both get 4-year sentences
  - They could have both had 3-years
- But it seems we should always defect and never cooperate
- No wonder Economics has been called the dismal science
Some PD examples

• There are lots of examples of the Prisoner’s Dilemma situations in the real world
• It makes it difficult for “players” to avoid the bad outcome of both defecting
  – Cheating on a cartel
  – Trade wars between countries
  – Arms races between countries
  – Advertising
  – Communal coffee pot
  – Class team project
**Cheating on a Cartel**

**Cartel**: association of firms with purpose of maintaining prices at a high level and restricting competition

– Cartel members' possible strategies range from abiding by their agreement to cheating
  i.e., can charge the cartel price or a lower one
– Cheating firms can increase profits
– The best strategy is charging the low price
Trade Wars Between Countries

- Free trade benefits both trading countries
- Tariffs can benefit one trading country
- Imposing tariffs can be a dominant strategy and establish a Nash equilibrium even though it may be inefficient
Advertising

- Advertising is expensive
- All firms advertising tends to equalize the effects
- Everyone would gain if no one advertised
- But firms increase their advertising to gain advantage
- Which makes their competition do the same
- It’s an arms race
Games Without Dominant Strategies

- In many games, players have no **dominant strategy**
- Player's strategy depends on others’ strategies
- If player's best strategy depends on another’s strategy, she has no dominant strategy

<table>
<thead>
<tr>
<th></th>
<th>Confess</th>
<th>Not Confess</th>
</tr>
</thead>
</table>
| **Ma**

  - Confess 6 years for Ma
  - 1 year for Pa 5 years for Ma

  - Not Confess 8 years for Ma
  - 0 years for Pa 4 years for Ma

  - Confess 5 years for Ma
  - 3 years for Pa

  - Not Confess 4 years for Ma
  - 2 years for Pa
Ma has no explicit dominant strategy, but there is a best one since Pa does have a dominant strategy (What is it?)
Pa’s Decision Tree

Pa does have a dominant strategy: confess
So Ma’s best strategy is to confess
Some games have no simple solution

Neither player has a dominant strategy.
There is no non-cooperative solution

Best strategy for each is to randomly choose 1 or 2
Repeated Games

• A repeated game is a game that the same players play more than once
• Repeated games differ from one-shot games since a player’s current actions can depend on the past behavior of other players
• Cooperation is encouraged
Payoff matrix for the generic two person dilemma game

<table>
<thead>
<tr>
<th></th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>(CC,CC)</td>
<td>(CD,DC)</td>
</tr>
<tr>
<td>Defect</td>
<td>(DC,CD)</td>
<td>(DD,DD)</td>
</tr>
</tbody>
</table>

- **(CC,CC)**: reward for mutual cooperation
- **(CD,DC)**: sucker’s payoff and temptation to defect
- **(DC,CD)**: temptation to defect and sucker’s payoff
- **(DD,DD)**: punishment for mutual defection

where C: cooperate and D: defect
Payoffs

• Four payoffs are involved
  – CC: Both players cooperate
  – CD: You cooperate, other defects (sucker’s payoff)
  – DC: You defect, other cooperates (temptation to defect)
  – DD: Both players defect

• Assigning values induces an ordering, with 24 possibilities (4!); three lead to “dilemma” games
  – **Prisoner’s dilemma**: DC > CC > DD > CD
  – **Chicken**: DC > CC > CD > DD
  – **Stag Hunt**: CC > DC > DD > CD
Chicken

- DC > CC > CD > DD
- Rebel without a cause scenario
- Two cars race toward one another
- Drivers choose to serve or not
  - Cooperation: swerving
  - Defecting: not swerving
- Optimal move: do exactly the opposite of other player
Stag Hunt

- CC > DC > DD > CD
- Two players on a stag hunt
- Hard task requiring coordination but with big shared payoff
- Hare seen, do you defect and chase it?
  - Cooperate: keep after the stag
  - Defect: switch to chasing hare
- Optimal play: do exactly what the other player(s) do
Prisoner’s dilemma

- DC > CC > DD > CD
- Optimal play: always defect
- Two rational players will always defect.
- Thus, (naïve) individual rationality subverts their common good
More examples of the PD in real life

• Communal coffeepot
  – Cooperate by making new pot of coffee if you take last cup
  – Defect by taking last cup and not making new pot, depending on the next coffee seeker to do it
  – DC > CC > DD > CD

• Class team project
  – Cooperate by doing your part well and on time
  – Defect by slacking, hoping other team members will come through and sharing benefits of good grade
  – (Arguable) DC > CC > DD > CD
Iterated Prisoner’s Dilemma

• Game theory: rational players should always defect when engaged in a PD situation
• In real situations, people don’t always do this
• Why not? Possible explanations:
  – People aren’t rational
  – Morality
  – Social pressure
  – Fear of consequences
  – Evolution of species-favoring genes
• Which make sense? How can we formalize?
Iterated Prisoner’s Dilemma

- **Key idea:** We often play more than one “game” with a given player.
- Players have knowledge of past games, including their choices and other players’ choices.
- Choice when playing against a player can be based on whether she’s cooperated in the past.
- Simulation first done by Robert Axelrod where programs played in a round-robin tournament (DC=5; CC=3; DD=1; CD=0).
- The simplest program won!
Some possible strategies

- Always defect
- Always cooperate
- Randomly choose
- Pavlovian (win-stay, lose-switch)
  Start always cooperate, switch to always defect when punished by other’s defection, switch back & forth on every punishment
- Tit-for-tat (TFT)
  Be nice, but punish defections: Start cooperating and, after that always do what other player did on previous round
- Joss
  Sneaky TFT that defects 10% of the time
- In an idealized (noise free) environment, TFT is both a very simple and very good strategy
Characteristics of Robust Strategies

Axelrod analyzed entries and identified characteristics

**Nice:** never defects first

**Provocable:** respond to defection by promptly defecting. Prompt response important; slow to anger a poor strategy; some programs tried even harder to take advantage

**Forgiving:** respond to single defections by defecting forever worked poorly. Better to respond to TIT with 0.9 TAT; might dampen echoes & prevent feuds

**Clear:** Clarity an important feature. With TFT you know what to expect and what will/won’t work. With too much randomness or bizarre strategies in program, competing programs cannot analyze and began to always defect.
Implications of Robust Strategies

• Succeed not by "beating" others, but by allowing both to do well. TFT never "wins" a single turn! It can't. It can never do better than tie (all C).

• You do well by motivating cooperative behavior from others ... the provocability part

• Envy is counterproductive. Doesn’t pay to get upset if someone does a few points better than you in a single encounter. To do well, others must also do well, e.g., business & its suppliers.
Implications of Robust Strategies

• Need not be smart to do well. TFT models cooperative relations with bacteria and hosts.
• Cosmic threats and promises aren’t necessary, though they may be helpful
• Central authority unnecessary, though it may be helpful
• Optimum strategy depends on environment. TFT not best program in all cases; too unforgiving of JOSS & too lenient with RANDOM
Emergence

• Process where larger entities, patterns, and regularities arise via interactions among smaller or simpler entities that themselves don’t exhibit such properties

• E.g.: Shape and behavior of a flock of birds or school of fish

• Might cooperation be an emergent property?
Required for emergent cooperation

• A non-zero sum situation
• Players equal in power; no discrimination or status differences
• Repeated encounters with others you can recognize
  Garages depending on repeat business versus those on busy highways. Being unlikely to ever see someone again => a non-iterated dilemma.
• Low temptation payoff
  If defecting makes you a billionaire, you're likely to do it. "Every person has a price"
Ecological model

• Assume ecological system supporting N players
• Players gain or loose points on each round
• After a round, worst players die, best multiply
• Environmental noise models that agent makes errors in following a strategy or misinterpret another’s choice
• A simple way of modeling this is described in The Computational Beauty of Nature
Evolutionary stable strategies

- Strategies do better or worse against other strategies
- Successful strategies should work well in a variety of environments
  - E.g.: ALL-C works well in a mono-culture of ALL-Cs but not in a mixed environment
- Successful strategies can “fight off mutations”
  - E.g.: ALL-D mono-culture is very resistant to invasions by any cooperating strategies
  - E.g.: TFT can be “invaded” by ALL-C
Population simulation

(a) TFT wins

(b) A noise free version with TFT winning

(c) 0.5% noise lets Pavlov win

Figure 17.3 Population simulations of the ecological version of the iterated Prisoner's Dilemma: (a) an idealized version that illustrates the rise of TFT; (b) a noise-free simulation with TFT winning; (c) with 0.5 percent noise Pavlov wins
If you are interested...

- **Axelrod Python**
  - [https://github.com/Axelrod-Python](https://github.com/Axelrod-Python)
  - Explore strategies for the Prisoners dilemma game
  - Over 100 strategies from the literature and some original ones
  - Run round robin tournaments with a variety of options
  - Population dynamics
- Easy to install
  - pip install axelrod
- Also includes notebooks
20th anniversary IPD competition (2004)

- New Tack Wins Prisoner's Dilemma
- Coordinating Team Players within a Noisy Iterated Prisoner’s Dilemma Tournament
- U. Southampton bot team won using covert channel to let Bots on the team recognize each other
- The 60 bots
  - Executed series of moves that signaled their ‘tribe’
  - Defect if other known to be outside tribe, coordinate if in tribe
  - Coordination was not just cooperation, but master/slave : defect/cooperate
Game Theory Relevance

• Game theory is important in more complex "games"
  – E.g.: multiplayer, non-zero-sum, complicated payoffs

• Repeated games add complexity to balance cooperation and competition

• Used in multi-agent systems and where agents form teams with humans
The worst resolution to the Valentine Prisoner’s Dilemma when YOU decide not to give your partner a present, but your PARTNER decides to testify against you in the armed robbery case.