Graduate AI

Lecture 19: Game Theory II

Teachers:
Zico Kolter
Ariel Procaccia (this time)
**Correlated equilibrium**

- Let $N = \{1, 2\}$ for simplicity.
- A mediator chooses a pair of strategies $(s_1, s_2)$ according to a distribution $p$ over $S^2$.
- Reveals $s_1$ to player 1 and $s_2$ to player 2.
- When player 1 gets $s_1 \in S$, he knows that the distribution over strategies of 2 is

$$
\Pr[s_2|s_1] = \frac{\Pr[s_1 \land s_2]}{\Pr[s_1]} = \frac{p(s_1, s_2)}{\sum_{s_2' \in S} p(s_1, s_2')}
$$
**Correlated equilibrium**

- Player 1 is best responding if for all $s'_1 \in S$
  \[
  \sum_{s_2 \in S} \Pr[s_2 | s_1] u_1(s_1, s_2) \geq \sum_{s_2 \in S} \Pr[s_2 | s_1] u_1(s'_1, s_2)
  \]

- Equivalently,
  \[
  \sum_{s_2 \in S} p(s_1, s_2) u_1(s_1, s_2) \geq \sum_{s_2 \in S} p(s_1, s_2) u_1(s'_1, s_2)
  \]

- $p$ is a **correlated equilibrium (CE)** if both players are best responding
GAME OF CHICKEN

http://youtu.be/u7hZ9jKrwvo
**Game of Chicken**

- **Social welfare** is the sum of utilities
- Pure NE: (C,D) and (D,C), social welfare = 5
- Mixed NE: both (1/2,1/2), social welfare = 4
- Optimal social welfare = 6

<table>
<thead>
<tr>
<th></th>
<th>Dare</th>
<th>Chicken</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dare</td>
<td>0,0</td>
<td>4,1</td>
</tr>
<tr>
<td>Chicken</td>
<td>1,4</td>
<td>3,3</td>
</tr>
</tbody>
</table>
GAME OF CHICKEN

• Correlated equilibrium:
  o (D,D): 0
  o (D,C): $\frac{1}{3}$
  o (C,D): $\frac{1}{3}$
  o (C,C): $\frac{1}{3}$

• Social welfare of CE = $\frac{16}{3}$
IMPLEMENTATION OF CE

• Instead of a mediator, use a hat!
• Balls in hat are labeled with “chicken” or “dare”, each blindfolded player takes a ball
• **Poll 1:** Which balls implement the distribution of slide 6?
  1. 1 chicken, 1 dare
  2. 2 chicken, 1 dare
  3. 2 chicken, 2 dare
  4. 3 chicken, 2 dare
CE vs. NE

• **Poll 2:** What is the relation between CE and NE?
  1. CE ⇒ NE
  2. NE ⇒ CE
  3. NE ⇔ CE
  4. NE ∥ CE

CE of slide 6 is NE?
CE As LP

- Can compute CE via linear programming in polynomial time!

\[
\text{find } \forall s_1, s_2 \in S, p(s_1, s_2) \\
\text{s.t. } \forall s_1, s'_1, s_2 \in S, \sum_{s_2 \in A} p(s_1, s_2)u_1(s_1, s_2) \geq \sum_{s_2 \in A} p(s_1, s_2)u_1(s'_1, s_2) \\
\forall s_1, s_2, s'_2 \in S, \sum_{s_1 \in A} p(s_1, s_2)u_2(s_1, s_2) \geq \sum_{s_1 \in A} p(s_1, s_2)u_2(s_1, s'_2) \\
\sum_{s_1, s_2 \in S} p(s_1, s_2) = 1 \\
\forall s_1, s_2 \in S, p(s_1, s_2) \in [0,1]
\]
A CURIOUS GAME

• Playing up is a dominant strategy for row player
• So column player would play left
• Therefore, (1,1) is the only Nash equilibrium outcome

\[
\begin{array}{cc}
1,1 & 3,0 \\
0,0 & 2,1 \\
\end{array}
\]
Commitment is Good

- Suppose the game is played as follows:
  - Row player commits to playing a row
  - Column player observes the commitment and chooses column

- Row player can commit to playing down!
COMMITMENT TO MIXED STRATEGY

• By committing to a mixed strategy, row player can guarantee a reward of 2.5

• Called a Stackelberg (mixed) strategy

\[
\begin{array}{c|cc}
& 0 & 1 \\
\hline
.49 & 1,1 & 3,0 \\
.51 & 0,0 & 2,1 \\
\end{array}
\]
Computing Stackelberg

• Theorem [Conitzer and Sandholm, EC 2006]: In 2-player normal form games, an optimal Stackelberg strategy can be found in poly time

• Theorem [ditto]: the problem is NP-hard when the number of players is $\geq 3$
TRACTABILITY: 2 PLAYERS

• For each pure follower strategy $s_2$, we compute via the LP below a strategy $x_1$ for the leader such that
  o Playing $s_2$ is a best response for the follower
  o Under this constraint, $x_1$ is optimal

• Choose $x_1^*$ that maximizes leader value

$$\max \sum_{s_1 \in S} x_1(s_1)u_1(s_1, s_2)$$

s.t. $\forall s'_2 \in S, \sum_{s_1 \in S} x_1(s_1)u_2(s_1, s_2) \geq \sum_{s_1 \in S} x_1(s_1)u_2(s_1, s'_2)$

$\sum_{s_1 \in S} x_1(s_1) = 1$

$\forall s_1 \in S, x_1(s_1) \in [0,1]$
APPLICATION: SECURITY

• Airport security: deployed at LAX
• Federal Air Marshals
• Coast Guard
• Idea:
  o Defender commits to mixed strategy
  o Attacker observes and best responds
**Security Games**

- Set of targets $T = \{1, \ldots, n\}$
- Set of $m$ security resources $\Omega$ available to the defender (leader)
- Set of schedules $\Sigma \subseteq 2^T$
- Resource $\omega$ can be assigned to one of the schedules in $A(\omega) \subseteq \Sigma$
- Attacker chooses one target to attack
SECURITY GAMES

- For each target $t$, there are four numbers: $\bar{u}_d(t) \geq \underline{u}_d(t)$, and $\bar{u}_a(t) \leq \underline{u}_a(t)$

- Randomized defender strategy induces coverage probabilities $c = (c_1, ..., c_n)$

- The utilities to the defender/attacker under $c$ if target $t$ is attacked are
  
  $u_d(t, c) = \bar{u}_d(t) \cdot c_t + \underline{u}_d(t)(1 - c_t)$
  
  $u_a(t, c) = \bar{u}_a(t) \cdot c_t + \underline{u}_a(t)(1 - c_t)$
This is a 2-player Stackelberg game, so we can compute an optimal strategy for the defender in polynomial time...?
The Element of Surprise

To help combat the terrorism threat, officials at Los Angeles International Airport are introducing a bold new idea into their arsenal: random security checkpoints. Can game theory help keep us safe?

WEB EXCLUSIVE

By Andrew Murr
Newsweek
Updated: 1:00 p.m. PT Sept 28, 2007

Sept. 28, 2007 - Security officials at Los Angeles International Airport now have a new weapon in their fight against terrorism: complete, baffling randomness. Anxious to thwart future terror attacks in the early stages while plotters are casing the airport, LAX security patrols have begun using a new software program called ARMOR. NEWSWEEK has learned, to make the placement of security checkpoints completely unpredictable. Now all airport security officials have to do is press a button labeled “Randomize,” and they can throw a sort of digital cloak of invisibility over where they place the cops’ antiterror checkpoints on any given day.
LIMITATIONS

• The defender knows the utility function of the attacker
  ○ Solution: machine learning

• The attacker perfectly observes the defender’s randomized strategy
  ○ MDPs, although this may not be a major concern

• The attacker is perfectly rational, i.e., best responds to the defender’s strategy
  ○ Solution: bounded rationality models
Testing bounded rationality

Game 1
Total: $1.4 = $1.5 - $0.1

Reward if successful
Penalty if caught by rangers
Money earned if successful

Next Game

[Kar et al., 2015]
**SUMMARY**

- **Terminology and algorithms:**
  - Correlated equilibrium: Polytime algorithm
  - Stackelberg game: Polytime algorithm
  - Security game

- **Nobel-prize-winning ideas:**
  - Correlated equilibrium 😊

- **Other big ideas:**
  - Stackelberg games for security