Computing Game-Theoretic Solutions

Vincent Conitzer
Duke University

Game theory

Multiple self-interested agents interacting in the same environment

What is an agent to do?

What is an agent to believe? (What are we to believe?)
Penalty kick example

Is this a "rational" outcome? If not, what is?
Multiagent systems

Goal: Blocked(Room0)

Goal: Clean(Room0)
Game playing
Real-world security applications

Airport security

- Where should checkpoints, canine units, etc. be deployed?
- Deployed at LAX airport and elsewhere

Milind Tambe’s TEAMCORE group (USC)

Federal Air Marshals

- Which flights get a FAM?

US Coast Guard

- Which patrol routes should be followed?
- Deployed in Boston, New York, Los Angeles
Mechanism design

Rating/voting systems

Auctions

Kidney exchanges

Prediction markets

overview: C., CACM March 2010
Outline

• Introduction to game theory (from CS/AI perspective)
  • Representing games
  • Standard solution concepts
    • (Iterated) dominance
    • Minimax strategies
    • Nash and correlated equilibrium
• Recent developments
  • Commitment: Stackelberg mixed strategies
  • Security applications
• Learning in games (time permitting)
  • Simple algorithms
  • Evolutionary game theory
  • Learning in Stackelberg games
Representing games
Rock-paper-scissors

<table>
<thead>
<tr>
<th></th>
<th>Column player aka. player 2 (simultaneously) chooses a column</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0, 0</td>
<td>-1, 1</td>
</tr>
<tr>
<td>-1, 1</td>
<td>0, 0</td>
</tr>
<tr>
<td>1, -1</td>
<td>-1, 1</td>
</tr>
<tr>
<td>-1, 1</td>
<td>1, -1</td>
</tr>
<tr>
<td>1, -1</td>
<td>0, 0</td>
</tr>
<tr>
<td>-1, 1</td>
<td>0, 0</td>
</tr>
<tr>
<td>1, -1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

Row player aka. player 1 chooses a row

A row or column is called an action or (pure) strategy

Row player’s utility is always listed first, column player’s second

Zero-sum game: the utilities in each entry sum to 0 (or a constant)
Three-player game would be a 3D table with 3 utilities per entry, etc.
Penalty kick
(also known as: matching pennies)

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>0,0</td>
<td>-1,1</td>
</tr>
<tr>
<td>R</td>
<td>-1,1</td>
<td>0,0</td>
</tr>
</tbody>
</table>
Security example

Terminal A

Terminal B

action

action
Security game

\[
\begin{array}{c|cc}
 & A & B \\
\hline
A & 0, 0 & -1, 2 \\
B & -1, 1 & 0, 0 \\
\end{array}
\]
“Chicken”

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

\[
\begin{array}{c|cc}
   & D & S \\
\hline
D & 0, 0 & -1, 1 \\
S & 1, -1 & -5, -5 \\
\end{array}
\]

not zero-sum
Modeling and representing games

### Normal-form games

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

### Extensive-form games

- Player 1 gets King, player 2 bet, player 1 check, player 2 call, player 2 fold.
- Player 1 gets Jack, player 2 bet, player 1 check, player 2 fold, player 2 fold.

### Bayesian games

- Players have different beliefs about the opponent.

### Stochastic games

- probabilistic outcomes in each decision point.

### Action-graph games

- graphical games

### MAIDs

- [Koller & Milch, IJCAI’01/GE’03]

### References

- [Kearns, Littman, Singh UAI’01]
- [Bhat & Leyton-Brown, UAI’04]
- [Jiang, Leyton-Brown, Bhat GEB’11]
A poker-like game

• Both players put 1 chip in the pot
• Player 1 gets a card (King is a winning card, Jack a losing card)
• Player 1 decides to raise (add one to the pot) or check
• Player 2 decides to call (match) or fold (P1 wins)
• If player 2 called, player 1’s card determines pot winner
Poker-like game in normal form

<table>
<thead>
<tr>
<th></th>
<th>cc</th>
<th>cf</th>
<th>fc</th>
<th>ff</th>
</tr>
</thead>
<tbody>
<tr>
<td>rr</td>
<td>0, 0</td>
<td>0, 0</td>
<td>1, -1</td>
<td>1, -1</td>
</tr>
<tr>
<td>rc</td>
<td>.5, -.5</td>
<td>1.5, -1.5</td>
<td>0, 0</td>
<td>1, -1</td>
</tr>
<tr>
<td>cr</td>
<td>-.5, .5</td>
<td>-.5, .5</td>
<td>1, -1</td>
<td>1, -1</td>
</tr>
<tr>
<td>cc</td>
<td>0, 0</td>
<td>1, -1</td>
<td>0, 0</td>
<td>1, -1</td>
</tr>
</tbody>
</table>
Our first solution concept: Dominance
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.
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 $i$”
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 (additional 2 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

<table>
<thead>
<tr>
<th></th>
<th>confess</th>
<th>don’t confess</th>
</tr>
</thead>
<tbody>
<tr>
<td>confess</td>
<td>-2, -2</td>
<td>0, -3</td>
</tr>
<tr>
<td>don’t confess</td>
<td>-3, 0</td>
<td>-1, -1</td>
</tr>
</tbody>
</table>
“Should I buy an SUV?”

purchasing (+gas, maintenance) cost

- cost: 5
- cost: 3

accident cost

- cost: 5
- cost: 8
- cost: 5

-10, -10  -7, -11
-11, -7  -8, -8
Back to the poker-like game

1 gets King
player 1
raise
check raise
1 gets Jack
player 1

“nature”

player 2
call
fold
call
fold
call
fold

player 2

2 1 1 1 -2 -1 1
1

<table>
<thead>
<tr>
<th></th>
<th>cc</th>
<th>cf</th>
<th>fc</th>
<th>ff</th>
</tr>
</thead>
<tbody>
<tr>
<td>rr</td>
<td>0, 0</td>
<td>0, 0</td>
<td>1, -1</td>
<td>1, -1</td>
</tr>
<tr>
<td>rc</td>
<td>.5, -.5</td>
<td>1.5, -1.5</td>
<td>0, 0</td>
<td>1, -1</td>
</tr>
<tr>
<td>cr</td>
<td>-.5, .5</td>
<td>-.5, .5</td>
<td>1, -1</td>
<td>1, -1</td>
</tr>
<tr>
<td>cc</td>
<td>0, 0</td>
<td>1, -1</td>
<td>0, 0</td>
<td>1, -1</td>
</tr>
</tbody>
</table>
Mixed strategies

- **Mixed strategy** for player $i = \text{probability distribution over player i’s (pure) strategies}
- E.g., $1/3, 1/3, 1/3$
- Example of dominance by a mixed strategy:

```
\begin{array}{ccc}
\frac{1}{2} & 3, 0 & 0, 0 \\
\frac{1}{2} & 0, 0 & 3, 0 \\
1, 0 & 1, 0 & \end{array}
```

Usage:
- $\sigma_i$ denotes a mixed strategy,
- $s_i$ denotes a pure strategy
Checking for dominance by mixed strategies

• Linear program for checking whether strategy $s_i^*$ is strictly dominated by a mixed strategy:
  • maximize $\varepsilon$
  • such that:
    – for any $s_{-i}$, $\sum_{s_i} p_{s_i} u_i(s_i, s_{-i}) \geq u_i(s_i^*, s_{-i}) + \varepsilon$
    – $\sum_{s_i} p_{s_i} = 1$

• Linear program for checking whether strategy $s_i^*$ is weakly dominated by a mixed strategy:
  • maximize $\sum_{s_{-i}} [(\sum_{s_i} p_{s_i} u_i(s_i, s_{-i})) - u_i(s_i^*, s_{-i})]$
  • such that:
    – for any $s_{-i}$, $\sum_{s_i} p_{s_i} u_i(s_i, s_{-i}) \geq u_i(s_i^*, s_{-i})$
    – $\sum_{s_i} p_{s_i} = 1$
Iterated dominance

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

<table>
<thead>
<tr>
<th></th>
<th>0, 0</th>
<th>1, -1</th>
<th>1, -1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1, 1</td>
<td>0, 0</td>
<td>-1, 1</td>
<td></td>
</tr>
<tr>
<td>-1, 1</td>
<td>1, -1</td>
<td>0, 0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0, 0</th>
<th>1, -1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1, 1</td>
<td>0, 0</td>
<td>0, 0</td>
</tr>
</tbody>
</table>
“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 solved

\[ \frac{2}{3} \times 100 \]

\[ \frac{2}{3} \times \frac{2}{3} \times 100 \]

\[ \ldots \]

0

\{ \text{dominated} \}

\{ \text{dominated after removal of (originally) dominated strategies} \}
Iterated dominance: path (in)dependence

Iterated weak dominance is **path-dependent**: sequence of eliminations may determine which solution we get (if any)
(whether or not dominance by mixed strategies allowed)
Leads to various NP-hardness results [Gilboa, Kalai, Zemel Math of OR '93; C. & Sandholm EC ’05, AAAI’05; Brandt, Brill, Fischer, Harrenstein TOCS ’11]

Iterated strict dominance is **path-independent**: elimination process will always terminate at the same point
(whether or not dominance by mixed strategies allowed)
Two computational questions for iterated dominance

• 1. Can a given strategy be eliminated using iterated dominance?

• 2. Is there some path of elimination by iterated dominance such that only one strategy per player remains?

• For strict dominance (with or without dominance by mixed strategies), both can be solved in polynomial time due to path-independence:
  – Check if any strategy is dominated, remove it, repeat

• For weak dominance, both questions are NP-hard (even when all utilities are 0 or 1), with or without dominance by mixed strategies [C., Sandholm 05]
  – Weaker version proved by [Gilboa, Kalai, Zemel 93]
Solving two-player zero-sum games
### How to play matching pennies

<table>
<thead>
<tr>
<th></th>
<th>Them</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Us</strong></td>
<td><strong>L</strong></td>
</tr>
<tr>
<td><strong>L</strong></td>
<td>1, -1</td>
</tr>
<tr>
<td><strong>R</strong></td>
<td>-1, 1</td>
</tr>
</tbody>
</table>

- Assume opponent knows our **mixed** strategy
- If we play L 60%, R 40%...
- … opponent will play R…
- … we get .6*(-1) + .4*(1) = -.2
- What’s optimal for us? What about rock-paper-scissors?
A locally popular sport

defend the 3

<table>
<thead>
<tr>
<th>go for 3</th>
<th>go for 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>-2, 2</td>
</tr>
<tr>
<td>-3, 3</td>
<td>0, 0</td>
</tr>
</tbody>
</table>
## Solving basketball

<table>
<thead>
<tr>
<th></th>
<th>Them</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Os</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0, 0</td>
</tr>
<tr>
<td></td>
<td>-3, 3</td>
</tr>
</tbody>
</table>

- If we 50% of the time defend the 3, opponent will shoot 3
  - We get \(0.5 \times (-3) + 0.5 \times 0 = -1.5\)
- Should defend the 3 more often: 60% of the time
- Opponent has choice between
  - Go for 3: gives them \(0.6 \times 0 + 0.4 \times 3 = 1.2\)
  - Go for 2: gives them \(0.6 \times 2 + 0.4 \times 0 = 1.2\)
- We get -1.2 (the **maximin** value)
Let’s change roles

<table>
<thead>
<tr>
<th></th>
<th>Them</th>
</tr>
</thead>
<tbody>
<tr>
<td>Us</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0, 0</td>
</tr>
<tr>
<td>2</td>
<td>-3, 3</td>
</tr>
</tbody>
</table>

- Suppose we know their strategy
- If 50% of the time they go for 3, then we defend 3
  - We get .5*(0)+.5*(-2) = -1
- Optimal for them: 40% of the time go for 3
  - If we defend 3, we get .4*(0)+.6*(-2) = -1.2
  - If we defend 2, we get .4*(-3)+.6*(0) = -1.2

von Neumann’s minimax theorem [1928]: maximin value = minimax value (~ linear programming duality)

This is the minimax value
Minimax theorem [von Neumann 1928]

• Maximin utility: \( \max_{\sigma_i} \min_{s_{-i}} u_i(\sigma_i, s_{-i}) \)
  \( = - \min_{\sigma_i} \max_{s_{-i}} u_{-i}(\sigma_i, s_{-i}) \)

• Minimax utility: \( \min_{\sigma_{-i}} \max_{s_i} u_i(s_i, \sigma_{-i}) \)
  \( = - \max_{\sigma_{-i}} \min_{s_i} u_{-i}(s_i, \sigma_{-i}) \)

• Minimax theorem:
  \( \max_{\sigma_i} \min_{s_{-i}} u_i(\sigma_i, s_{-i}) = \min_{\sigma_{-i}} \max_{s_i} u_i(s_i, \sigma_{-i}) \)

• Minimax theorem does not hold with pure strategies only (example?)
Practice games

<table>
<thead>
<tr>
<th></th>
<th>20, -20</th>
<th>0, 0</th>
<th>10, -10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>0, 0</td>
<td>10, -10</td>
<td>8, -8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>20, -20</th>
<th>0, 0</th>
<th>10, -10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>0, 0</td>
<td>10, -10</td>
<td>8, -8</td>
</tr>
</tbody>
</table>
Back to the poker-like game, again

- To make player 1 indifferent between bb and bs, we need:
  utility for bb = 0*P(cc)+1*(1-P(cc)) = .5*P(cc)+0*(1-P(cc)) = utility for bs
  That is, P(cc) = 2/3

- To make player 2 indifferent between cc and fc, we need:
  utility for cc = 0*P(bb)+(-.5)*(1-P(bb)) = -1*P(bb)+0*(1-P(bb)) = utility for fc
  That is, P(bb) = 1/3
A brief history of the minimax theorem

Computing minimax strategies

- maximize $v_R$
  
  subject to

  for all $c$, $\sum_r p_r u_R(r, c) \geq v_R$

  $\sum_r p_r = 1$
Equilibrium notions for general-sum games
General-sum games

- You could still play a minimax strategy in general-sum games
  - I.e., pretend that the opponent is only trying to hurt you

- But this is not rational:

<table>
<thead>
<tr>
<th>0, 0</th>
<th>3, 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 0</td>
<td>2, 1</td>
</tr>
</tbody>
</table>

- If Column was trying to hurt Row, Column would play Left, so Row should play Down
- In reality, Column will play Right (strictly dominant), so Row should play Up
- Is there a better generalization of minimax strategies in zero-sum games to general-sum games?
Nash equilibrium [Nash 1950]

• A profile (= strategy for each player) so that no player wants to deviate

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0, 0</td>
<td>-1, 1</td>
</tr>
<tr>
<td>S</td>
<td>1, -1</td>
<td>-5, -5</td>
</tr>
</tbody>
</table>

• This game has another Nash equilibrium in mixed strategies – both play D with 80%
Nash equilibria of “chicken”…

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0, 0</td>
<td>1, -1</td>
</tr>
<tr>
<td>S</td>
<td>1, -1</td>
<td>-5, -5</td>
</tr>
</tbody>
</table>

- 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.
The presentation game

<table>
<thead>
<tr>
<th>Pay attention (A)</th>
<th>Do not pay attention (NA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Put effort into presentation (E)</td>
<td>2, 2</td>
</tr>
<tr>
<td>Do not put effort into presentation (NE)</td>
<td>-7, -8</td>
</tr>
</tbody>
</table>

- Pure-strategy Nash equilibria: (E, A), (NE, NA)
- Mixed-strategy Nash equilibrium:

  \[ \left( \frac{4}{5} \text{ E, } \frac{1}{5} \text{ NE}, \frac{1}{10} \text{ A, } \frac{9}{10} \text{ NA} \right) \]

  - Utility -7/10 for presenter, 0 for audience
The “equilibrium selection problem”

• You are about to play a game that you have never played before with a person that you have never met
• According to which equilibrium should you play?
• Possible answers:
  – Equilibrium that maximizes the sum of utilities (social welfare)
  – Or, at least not a Pareto-dominated equilibrium
  – So-called focal equilibria
    • “Meet in Paris” game: You and a friend were supposed to meet in Paris at noon on Sunday, but you forgot to discuss where and you cannot communicate. All you care about is meeting your friend. Where will you go?
  – Equilibrium that is the convergence point of some learning process
  – An equilibrium that is easy to compute
  – …
• Equilibrium selection is a difficult problem
Computing a single Nash equilibrium

“Together with factoring, the complexity of finding a Nash equilibrium is in my opinion the most important concrete open question on the boundary of $P$ today.”

Christos Papadimitriou, *STOC’01*

- PPAD-complete to compute one Nash equilibrium in a two-player game [Daskalakis, Goldberg, Papadimitriou *STOC’06 / SIAM J. Comp. ‘09; Chen & Deng FOCS’06 / Chen, Deng, Teng JACM’09]

- Is one Nash equilibrium all we need to know?
A useful reduction (SAT $\rightarrow$ game)

[C. & Sandholm IJCAI’03, Games and Economic Behavior ‘08]
(Earlier reduction with weaker implications: Gilboa & Zemel GEB ‘89)

Formula: $(x_1 \lor -x_2)$ and $(-x_1 \lor x_2)$

Solutions:
- $x_1 = \text{true}, x_2 = \text{true}$
- $x_1 = \text{false}, x_2 = \text{false}$

Game:

<table>
<thead>
<tr>
<th></th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$+x_1$</th>
<th>$-x_1$</th>
<th>$+x_2$</th>
<th>$-x_2$</th>
<th>$(x_1 \lor -x_2)$</th>
<th>$(-x_1 \lor x_2)$</th>
<th>default</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>-2,-2</td>
<td>-2,-2</td>
<td>0,-2</td>
<td>0,-2</td>
<td>2,-2</td>
<td>2,-2</td>
<td>-2,-2</td>
<td>-2,-2</td>
<td>0,1</td>
</tr>
<tr>
<td>$x_2$</td>
<td>-2,-2</td>
<td>-2,-2</td>
<td>2,-2</td>
<td>2,-2</td>
<td>0,-2</td>
<td>0,-2</td>
<td>-2,-2</td>
<td>-2,-2</td>
<td>0,1</td>
</tr>
<tr>
<td>$+x_1$</td>
<td>-2,0</td>
<td>-2,2</td>
<td>1,1</td>
<td>-2,-2</td>
<td>1,1</td>
<td>1,1</td>
<td>-2,0</td>
<td>-2,2</td>
<td>0,1</td>
</tr>
<tr>
<td>$-x_1$</td>
<td>-2,0</td>
<td>-2,2</td>
<td>-2,-2</td>
<td>1,1</td>
<td>1,1</td>
<td>1,1</td>
<td>-2,2</td>
<td>-2,0</td>
<td>0,1</td>
</tr>
<tr>
<td>$+x_2$</td>
<td>-2,2</td>
<td>-2,0</td>
<td>1,1</td>
<td>1,1</td>
<td>-2,-2</td>
<td>1,1</td>
<td>-2,0</td>
<td>-2,2</td>
<td>0,1</td>
</tr>
<tr>
<td>$-x_2$</td>
<td>-2,2</td>
<td>-2,0</td>
<td>1,1</td>
<td>1,1</td>
<td>-2,-2</td>
<td>1,1</td>
<td>-2,0</td>
<td>-2,2</td>
<td>0,1</td>
</tr>
</tbody>
</table>

- Every satisfying assignment (if there are any) corresponds to an equilibrium with utilities 1, 1; exactly one additional equilibrium with utilities $\epsilon, \epsilon$ that always exists
- Evolutionarily stable strategies $\Sigma_2$ $\text{P}$-complete [C. WINE 2013]
Some algorithm families for computing Nash equilibria of 2-player normal-form games

- for both $i$, for any $s_i \in S_i - X_i$, $p_i(s_i) = 0$
- for both $i$, for any $s_i \in X_i$, $\sum p_i(s_i)u_i(s_i, s_{-i}) = u_i$
- for both $i$, for any $s_i \in S_i - X_i$, $\sum p_i(s_i)u_i(s_i, s_{-i}) \leq u_i$

Search over supports / MIP

[Dickhaut & Kaplan, Mathematica J. ‘91]
[Porter, Nudelman, Shoham AAAI’04 / GEB’08]
[Sandholm, Gilpin, C. AAAI’05]

Special cases / subroutines
[C & Sandholm AAAI’05, AAMAS’06; Benisch, Davis, Sandholm AAAI’06 / JAIR’10; Kontogiannis & Spirakis APPROX’11; Adsul, Garg, Mehta, Sohoni STOC’11; …]

Approximate equilibria
[Brown ’51 / C. ’09 / Goldberg, Savani, Sørensen, Ventre ’11; Althöfer ’94, Lipton, Markakis, Mehta ‘03, Daskalakis, Mehta, Papadimitriou ‘06, ’07, Feder, Nazerzadeh, Saberi ’07, Tsaknakis & Spirakis ‘07, Spirakis ‘08, Bosse, Byrka, Markakis ‘07, …]
Search-based approaches (for 2 players)

• Suppose we know the support $X_i$ of each player $i$’s mixed strategy in equilibrium
  – That is, which pure strategies receive positive probability

• Then, we have a linear feasibility problem:
  – for both $i$, for any $s_i \in S_i - X_i$, $p_i(s_i) = 0$
  – for both $i$, for any $s_i \in X_i$, $\sum p_{-i}(s_{-i})u_i(s_i, s_{-i}) = u_i$
  – for both $i$, for any $s_i \in S_i - X_i$, $\sum p_{-i}(s_{-i})u_i(s_i, s_{-i}) \leq u_i$

• Thus, we can search over possible supports
  – This is the basic idea underlying methods in [Dickhaut & Kaplan 91; Porter, Nudelman, Shoham AAAI04/GEB08]

• Dominated strategies can be eliminated
Solving for a Nash equilibrium using MIP (2 players)

[Sandholm, Gilpin, C. AAAI’05]

• maximize \textit{whatever you like} (e.g., social welfare)

• subject to

  – for both \(i\), for any \(s_i\), \(\Sigma_{s_{-i}} p_{s_{-i}} u_i(s_i, s_{-i}) = u_{s_i}\)
  
  – for both \(i\), for any \(s_i\), \(u_i \geq u_{s_i}\)
  
  – for both \(i\), for any \(s_i\), \(p_{s_i} \leq b_{s_i}\)
  
  – for both \(i\), for any \(s_i\), \(u_i - u_{s_i} \leq M(1 - b_{s_i})\)
  
  – for both \(i\), \(\Sigma s_{s_i} p_{s_i} = 1\)

• \(b_{s_i}\) is a binary variable indicating whether \(s_i\) is in the support, \(M\) is a large number
Lemke-Howson algorithm (1-slide sketch!)

<table>
<thead>
<tr>
<th></th>
<th>GREEN</th>
<th>ORANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>1, 0</td>
<td>0, 1</td>
</tr>
<tr>
<td>BLUE</td>
<td>0, 2</td>
<td>1, 0</td>
</tr>
</tbody>
</table>

- Strategy profile = pair of points
- Profile is an equilibrium iff every pure strategy is either a best response or unplayed
- I.e. equilibrium = pair of points that includes all the colors
  - ... except, pair of bottom points doesn’t count (the “artificial equilibrium”)
- Walk in some direction from the artificial equilibrium; at each step, throw out the color used twice
Correlated equilibrium [Aumann ‘74]

<table>
<thead>
<tr>
<th></th>
<th>0, 0</th>
<th>0, 1</th>
<th>1, 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1/6</td>
<td>1/6</td>
</tr>
<tr>
<td>1/6</td>
<td>0</td>
<td>0</td>
<td>1/6</td>
</tr>
<tr>
<td>0, 1</td>
<td>1/6</td>
<td>1/6</td>
<td>0</td>
</tr>
<tr>
<td>1/6</td>
<td>1/6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Correlated equilibrium LP

maximize \textit{whatever}

subject to

for all $r$ and $r'$, $\sum_c p_{r,c} u_R(r, c) \geq \sum_c p_{r,c} u_R(r', c)$  
Row incentive constraint

for all $c$ and $c'$, $\sum_r p_{r,c} u_C(r, c) \geq \sum_r p_{r,c} u_C(r, c')$  
Column incentive constraint

$\sum_{r,c} p_{r,c} = 1$  
distributional constraint
Recent developments
Questions raised by security games

- Equilibrium selection?

- How should we model temporal / information structure?

- What structure should utility functions have?

- Do our algorithms scale?
Observing the defender’s distribution in security

This model is not uncontroversial... [Pita, Jain, Tambe, Ordóñez, Kraus AIJ’10; Korzhyk, Yin, Kiekintveld, C., Tambe JAIR’11; Korzhyk, C., Parr AAMAS’11]
Commitment (Stackelberg strategies)
Commitment

Unique Nash equilibrium (iterated strict dominance solution)

- Suppose the game is played as follows:
  - Player 1 commits to playing one of the rows,
  - Player 2 observes the commitment and then chooses a column

- Optimal strategy for player 1: commit to Down
Commitment as an extensive-form game

• For the case of committing to a pure strategy:

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Player 2</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>Right</td>
<td>Right</td>
</tr>
<tr>
<td>Down</td>
<td>Left</td>
<td>Left</td>
</tr>
</tbody>
</table>

1, 1  3, 0  0, 0  2, 1
Commitment to mixed strategies

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>.49</td>
<td>1, 1</td>
<td>3, 0</td>
</tr>
<tr>
<td>.51</td>
<td>0, 0</td>
<td>2, 1</td>
</tr>
</tbody>
</table>

Sometimes also called a **Stackelberg (mixed) strategy**
Commitment as an extensive-form game…

• … for the case of committing to a mixed strategy:

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 0) = Up</td>
<td>Left (1, 1)</td>
</tr>
<tr>
<td>(.5, .5)</td>
<td>Right (3, 0)</td>
</tr>
<tr>
<td>(0, 1) = Down</td>
<td>Left (.5, .5)</td>
</tr>
<tr>
<td>Right (2.5, .5)</td>
<td>Right (0, 0)</td>
</tr>
</tbody>
</table>

• Economist: Just an extensive-form game, nothing new here

• Computer scientist: Infinite-size game! Representation matters
Computing the optimal mixed strategy to commit to

[C. & Sandholm EC’06, von Stengel & Zamir GEB’10]

• Separate LP for every column $c^*$:

\[
\begin{align*}
\text{maximize} & \quad \sum_r p_r u_R(r, c^*) \\
\text{subject to} & \quad \sum_c \sum_r p_r u_C(r, c) \geq \sum_r p_r u_C(r, c) \\
& \quad \sum_r p_r = 1
\end{align*}
\]

Row utility

Column optimality

distributional constraint
On the game we saw before

<table>
<thead>
<tr>
<th></th>
<th>1, 1</th>
<th>3, 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>0, 0</td>
<td>2, 1</td>
</tr>
</tbody>
</table>

maximize $1x + 0y$

subject to

$1x + 0y \geq 0x + 1y$

$x + y = 1$

$x \geq 0$

$y \geq 0$

maximize $3x + 2y$

subject to

$0x + 1y \geq 1x + 0y$

$x + y = 1$

$x \geq 0$

$y \geq 0$
Visualization

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>C</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>0,1</td>
<td>1,0</td>
<td>0,0</td>
</tr>
<tr>
<td>M</td>
<td>4,0</td>
<td>0,1</td>
<td>0,0</td>
</tr>
<tr>
<td>D</td>
<td>0,0</td>
<td>1,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>

(0,1,0) = M
(1,0,0) = U
(0,0,1) = D
Generalizing beyond zero-sum games

Minimax, Nash, Stackelberg all agree in zero-sum games

zero-sum games

minimax strategies

Nash equilibrium

Stackelberg mixed strategies
Other nice properties of commitment to mixed strategies

- No equilibrium selection problem

- Leader’s payoff at least as good as any Nash eq. or even correlated eq.

\[
\begin{array}{cc}
0, 0 & -1, 1 \\
1, -1 & -5, -5 \\
\end{array}
\]

(von Stengel & Zamir [GEB ‘10]; see also C. & Korzhyk [AAAI ‘11], Letchford, Korzhyk, C. [JAAMAS’14])

More discussion: V. Conitzer. Should Stackelberg Mixed Strategies Be Considered a Separate Solution Concept? [LOFT 2014]
Committing to a correlated strategy [C. & Korzhyk AAAI’11]

<table>
<thead>
<tr>
<th></th>
<th>1, 1</th>
<th>3, 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>.4</td>
<td>.2</td>
</tr>
<tr>
<td>2, 1</td>
<td>.1</td>
<td>.3</td>
</tr>
</tbody>
</table>
LP for optimal correlated strategy to commit to

maximize $\sum_{r,c} p_{r,c} u_C(r, c)$ \quad \text{leader utility}

subject to

for all $c$ and $c'$, $\sum_r p_{r,c} u_C(r, c) \geq \sum_r p_{r,c} u_C(r, c')$ \quad \text{Column incentive constraint}

$\sum_{r,c} p_{r,c} = 1$ \quad \text{distributional constraint}
Proposition 1. There exists an optimal correlated strategy to commit to in which the follower always gets the same recommendation.
3-player example

Leader

Unique optimal correlated strategy to commit to:

50%

50%

Different from Stackelberg / CE
Stackelberg mixed strategies deserve recognition as a separate solution concept!

- Seeing it only as a solution of a modified (extensive-form) game makes it hard to see...
  - when it coincides with other solution concepts
  - how utilities compare to other solution concepts
  - how to compute solutions
  - ...

- Does not mean it’s not also useful to think of it as a backward induction solution

- Similar story for correlated equilibrium
Some other work on commitment in unrestricted games

<table>
<thead>
<tr>
<th></th>
<th>2, 2</th>
<th>-1, 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>-7, -8</td>
<td>0, 0</td>
<td></td>
</tr>
</tbody>
</table>

**Normal-form games**

Learning to commit [Letchford, C., Munagala SAGT’09]

Correlated strategies [C. & Korzhyk AAAI’11]

Uncertain observability [Korzhyk, C., Parr AAMAS’11]

**Extensive-form games**

[Ordóñez, Kraus AAMAS’08; Letchford, C., Munagala SAGT’09; Pita, Jain, Tambe, Ordóñez, Kraus AIJ’10; Jain, Kiekintveld, Tambe AAMAS’11; …]

**Commitment in Bayesian games**

[C. & Sandholm EC’06; Paruchuri, Pearce, Mareckl, Tambe, Ordóñez, Kraus AAMAS’08; Letchford, C., Munagala SAGT’09; Pita, Jain, Tambe, Ordóñez, Kraus AIJ’10; Jain, Kiekintveld, Tambe AAMAS’11; …]

**Stochastic games**

[Letchford, MacDermed, C., Parr, Isbell, AAAI’12]
Security games
Example security game

- 3 airport terminals to defend (A, B, C)
- Defender can place checkpoints at 2 of them
- Attacker can attack any 1 terminal

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>{A, B}</td>
<td>0, -1</td>
<td>0, -1</td>
<td>-2, 3</td>
</tr>
<tr>
<td>{A, C}</td>
<td>0, -1</td>
<td>-1, 1</td>
<td>0, 0</td>
</tr>
<tr>
<td>{B, C}</td>
<td>-1, 1</td>
<td>0, -1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>
Security resource allocation games
[Kiekintveld, Jain, Tsai, Pita, Ordóñez, Tambe AAMAS’09]

• Set of targets $T$
• Set of security resources $\Omega$ available to the defender (leader)
• Set of schedules $S \subseteq 2^T$
• Resource $\omega$ can be assigned to one of the schedules in $A(\omega) \subseteq S$
• Attacker (follower) chooses one target to attack
• Utilities: $U_d^c(t), U_a^c(t)$ if the attacked target is defended, $U_d^u(t), U_a^u(t)$ otherwise
• $U_d^c(t) \geq U_d^u(t); U_a^c(t) \leq U_a^u(t)$
Game-theoretic properties of security resource allocation games [Korzhyk, Yin, Kiekintveld, C., Tambe JAIR’11]

- For the defender:
  Stackelberg strategies are also Nash strategies
  - minor assumption needed
  - not true with multiple attacks

- Interchangeability property for Nash equilibria (“solvable”)
  - no equilibrium selection problem
  - still true with multiple attacks

\[
\begin{array}{ccc}
1, 2 & 1, 0 & 2, 2 \\
1, 1 & 1, 0 & 2, 1 \\
0, 1 & 0, 0 & 0, 1 \\
\end{array}
\]
Scalability in security games

**basic model**

[Kiekintveld, Jain, Tsai, Pita, Ordóñez, Tambe AAMAS’09; Korzhyk, C., Parr, AAAI’10; Jain, Kardeș, Kiekintveld, Ordóñez, Tambe AAAI’10; Korzhyk, C., Parr, IJCAI’11]

**compact linear/integer programs**

Maximize $U^d(t^*) \sum_{s_1 \in S_{t^*}} c_{\omega,s} + U^a(t) \left(1 - \sum_{s_1 \in S_{t^*}} c_{\omega,s}\right)$

Subject to

$\forall \omega: \sum_{s} c_{\omega,s} \leq 1$

$\forall t: \sum_{s} \sum_{s_1 \in S_t} c_{\omega,s} \leq 1$

$\forall t: U^a(t) \sum_{s} \sum_{s_1 \in S_t} c_{\omega,s} + U^u(t) \left(1 - \sum_{s} \sum_{s_1 \in S_t} c_{\omega,s}\right)$

$\leq U^a(t^*) \sum_{s} \sum_{s_1 \in S_{t^*}} c_{\omega,s} + U^u(t^*) \left(1 - \sum_{s} \sum_{s_1 \in S_{t^*}} c_{\omega,s}\right)$

Defender utility

Marginal probability of $t^*$ being defended (~)

Distributional constraints

Attacker optimality

**Strategy generation**

$\min \ \ u$

subject to

$\sigma_h(s_{h0}) + \sigma_h(s_{h1}) + \ldots \sigma_h(s_{hk}) = 1$

$u \geq \sigma_h(s_{h0}) \cdot u(s_{h0}, s_{h0}) + \sigma_h(s_{h1}) \cdot u(s_{h1}, s_{h0}) + \ldots \sigma_h(s_{hk}) \cdot u(s_{hk}, s_{h0})$

$u \geq \sigma_h(s_{h0}) \cdot u(s_{h0}, s_{h1}) + \sigma_h(s_{h1}) \cdot u(s_{h1}, s_{h1}) + \ldots \sigma_h(s_{hk}) \cdot u(s_{hk}, s_{hk})$

$\ldots$

$u \geq \sigma_h(s_{h0}) \cdot u(s_{h0}, s_{hk}) + \sigma_h(s_{h1}) \cdot u(s_{h1}, s_{hk}) + \ldots \sigma_h(s_{hk}) \cdot u(s_{hk}, s_{hk})$

$\sigma_h(s_{h0}) \cdot u(s_{h0}, s_{h0}) + \sigma_h(s_{h1}) \cdot u(s_{h1}, s_{h0}) + \ldots \sigma_h(s_{hk}) \cdot u(s_{hk}, s_{h0})$

$\sigma_h(s_{h0}) \cdot u(s_{h0}, s_{h1}) + \sigma_h(s_{h1}) \cdot u(s_{h1}, s_{h1}) + \ldots \sigma_h(s_{hk}) \cdot u(s_{hk}, s_{hk})$

$\ldots$

$\sigma_h(s_{h0}) \cdot u(s_{h0}, s_{hk}) + \sigma_h(s_{h1}) \cdot u(s_{h1}, s_{hk}) + \ldots \sigma_h(s_{hk}) \cdot u(s_{hk}, s_{hk})$
Compact LP

- Cf. ERASER-C algorithm by Kiekintveld et al. [2009]
- Separate LP for every possible $t^*$ attacked:

Maximize $U^c_d(t^*) \sum_{\omega} \sum_{s: t^* \in s} c_{\omega,s} + U^u_d(t^*) \left(1 - \sum_{\omega} \sum_{s: t^* \in s} c_{\omega,s}\right)$

Subject to

$\forall \omega: \sum_{s} c_{\omega,s} \leq 1$

$\forall t: \sum_{\omega} \sum_{s: t \in s} c_{\omega,s} \leq 1$

$\forall t: U^c_a(t) \sum_{\omega} \sum_{s: t \in s} c_{\omega,s} + U^u_a(t) \left(1 - \sum_{\omega} \sum_{s: t \in s} c_{\omega,s}\right) \leq U^c_a(t^*) \sum_{\omega} \sum_{s: t^* \in s} c_{\omega,s} + U^u_a(t^*) \left(1 - \sum_{\omega} \sum_{s: t^* \in s} c_{\omega,s}\right)$

Defender utility

Marginal probability of $t^*$ being defended (?)

Distributional constraints

Attacker optimality
Counter-example to the compact LP

- LP suggests that we can cover every target with probability 1...
- … but in fact we can cover at most 3 targets at a time
Birkhoff-von Neumann theorem

- Every *doubly stochastic* $n \times n$ matrix can be represented as a convex combination of $n \times n$ permutation matrices

\[
\begin{bmatrix}
.1 & .4 & .5 \\
.3 & .5 & .2 \\
.6 & .1 & .3
\end{bmatrix}
\]

\[
\begin{array}{ccc}
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0
\end{array} + .1 \begin{array}{ccc}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{array} + .5 \begin{array}{ccc}
0 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 0
\end{array} + .3 \begin{array}{ccc}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{array}
\]

- Decomposition can be found in polynomial time $O(n^{4.5})$, and the size is $O(n^2)$ [Dulmage and Halperin, 1955]

- Can be extended to *rectangular* doubly *substochastic* matrices
Schedules of size 1 using BvN

\[ \omega_1 \]
\[ \omega_2 \]

\[ t_1 \]
\[ t_2 \]
\[ t_3 \]

\[
\begin{array}{ccc}
0 & 0 & 1 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{array}
\]

\[
\begin{array}{ccc}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{array}
\]

\[
\begin{array}{ccc}
1 & 0 & 0 \\
0 & 0 & 1 \\
\end{array}
\]
## Algorithms & complexity

[Korzhyk, C., Parr AAAI’10]

<table>
<thead>
<tr>
<th>Schedules</th>
<th>Homogeneous Resources</th>
<th>Heterogeneous resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size 1</td>
<td>P</td>
<td>P (BvN theorem)</td>
</tr>
<tr>
<td>Size ≤2, bipartite</td>
<td>P (BvN theorem)</td>
<td>NP-hard (SAT)</td>
</tr>
<tr>
<td>Size ≤2</td>
<td>P (constraint generation)</td>
<td>NP-hard</td>
</tr>
<tr>
<td>Size ≥3</td>
<td>NP-hard (3-COVER)</td>
<td>NP-hard</td>
</tr>
</tbody>
</table>

Also: security games on graphs

[Letchford, C. AAAI’13]
Security games with multiple attacks

[Korzhyk, Yin, Kiekintveld, C., Tambe JAIR’11]

- The attacker can choose multiple targets to attack

  ![Diagram showing multiple targets and attacks]

- The utilities are added over all attacked targets

- Stackelberg NP-hard; Nash polytime-solvable and interchangeable [Korzhyk, C., Parr IJCAI‘11]

  - Algorithm generalizes ORIGAMI algorithm for single attack
    [Kiekintveld, Jain, Tsai, Pita, Ordóñez, Tambe AAMAS’09]
Actual Security Schedules: Before vs. After
Boston, Coast Guard – “PROTECT” algorithm

slide courtesy of Milind Tambe

Before PROTECT

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After PROTECT

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Industry port partners comment:
“The Coast Guard seems to be everywhere, all the time.”
Data from LAX checkpoints before and after “ARMOR” algorithm

not a controlled experiment!

Firearm Violations | Drug Related Offenses | Miscellaneous | Total
---|---|---|---
(pre) 4/17/06 to 7/31/07
1/1/08 to 12/31/08
1/1/09 to 12/31/09
1/1/10 to 12/31/10
0 | 0 | 0 | 0
10 | 20 | 30 | 40
20 | 30 | 40 | 50
30 | 40 | 50 | 60
40 | 50 | 60 | 70
50 | 60 | 70 | 80
60 | 70 | 80 | 90
70 | 80 | 90 | 100
80 | 90 | 100 | 110
90 | 100 | 110 | 120
100 | 110 | 120 | 130
110 | 120 | 130 | 140
Placing checkpoints in a city

[Tsai, Yin, Kwak, Kempe, Kiekintveld, Tambe AAAI’10; Jain, Korzhyk, Vaněk, C., Pěchouček, Tambe AAMAS’11; Jain, C., Tambe AAMAS’13]
Learning in games
Learning in (normal-form) games

• Learn how to play a game by
  – playing it many times, and
  – updating your strategy based on experience

• Why?
  – Some of the game’s utilities (especially the other players’) may be unknown to you
  – The other players may not be playing an equilibrium strategy
  – Computing an optimal strategy can be hard
  – Learning is what humans typically do
  – …

• Does learning converge to equilibrium?
Iterated best response

- In the first round, play something arbitrary
- In each following round, play a best response against what the other players played in the previous round
- If all players play this, it can converge (i.e., we reach an equilibrium) or cycle

\[
\begin{array}{ccc}
0, 0 & -1, 1 & 1, -1 \\
1, -1 & 0, 0 & -1, 1 \\
-1, 1 & 1, -1 & 0, 0 \\
\end{array}
\]

\[
\begin{array}{cc}
-1, -1 & 0, 0 \\
0, 0 & -1, -1 \\
\end{array}
\]

*rock-paper-scissors*

- **Alternating best response**: players alternatingly change strategies: one player best-responds each odd round, the other best-responds each even round
Fictitious play [Brown 1951]

- In the first round, play something arbitrary
- In each following round, play a best response against the empirical distribution of the other players’ play
  - I.e., as if other player randomly selects from his past actions
- Again, if this converges, we have a Nash equilibrium
- Can still fail to converge...

\[
\begin{array}{ccc}
0, 0 & -1, 1 & 1, -1 \\
1, -1 & 0, 0 & -1, 1 \\
-1, 1 & 1, -1 & 0, 0 \\
\end{array}
\]

\[
\begin{array}{cc}
-1, -1 & 0, 0 \\
0, 0 & -1, -1 \\
\end{array}
\]

*rock-paper-scissors*
Fictitious play on rock-paper-scissors

<table>
<thead>
<tr>
<th></th>
<th>Row</th>
<th>Column</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td><img src="rock" alt="" /></td>
<td><img src="paper" alt="" /></td>
</tr>
<tr>
<td>-1, 1</td>
<td><img src="scissors" alt="" /></td>
<td><img src="scissors" alt="" /></td>
</tr>
<tr>
<td>1, -1</td>
<td><img src="scissors" alt="" /></td>
<td><img src="rock" alt="" /></td>
</tr>
<tr>
<td>1, -1</td>
<td><img src="paper" alt="" /></td>
<td><img src="paper" alt="" /></td>
</tr>
<tr>
<td>-1, 1</td>
<td><img src="scissors" alt="" /></td>
<td><img src="scissors" alt="" /></td>
</tr>
<tr>
<td>-1, 1</td>
<td><img src="scissors" alt="" /></td>
<td><img src="scissors" alt="" /></td>
</tr>
</tbody>
</table>

30% R, 50% P, 20% S  
30% R, 20% P, 50% S
Does the empirical distribution of play converge to equilibrium?

• ... for iterated best response?
• ... for fictitious play?

<table>
<thead>
<tr>
<th></th>
<th>3, 0</th>
<th>1, 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2</td>
<td>2, 1</td>
<td></td>
</tr>
</tbody>
</table>
Fictitious play is guaranteed to converge in...

- Two-player zero-sum games [Robinson 1951]
- Generic 2x2 games [Miyasawa 1961]
- Games solvable by iterated strict dominance [Nachbar 1990]
- Weighted potential games [Monderer & Shapley 1996]
- Not in general [Shapley 1964]
- But, fictitious play always converges to the set of $\frac{1}{2}$-approximate equilibria [C. 2009; more detailed analysis by Goldberg, Savani, Sørensen, Ventre 2011]
Shapley’s game on which fictitious play does not converge

starting with \((U, C)\):

<table>
<thead>
<tr>
<th></th>
<th>0, 0</th>
<th>0, 1</th>
<th>1, 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>1, 0</td>
<td>0, 0</td>
<td>0, 1</td>
</tr>
<tr>
<td>1, 0</td>
<td>0, 0</td>
<td>0, 1</td>
<td>0, 0</td>
</tr>
<tr>
<td>0, 1</td>
<td>1, 0</td>
<td>0, 0</td>
<td>0, 0</td>
</tr>
</tbody>
</table>
“Teaching”

- Suppose you are playing against a player that uses one of these learning strategies
  - Fictitious play, anything with no regret, …
- Also suppose you are very patient, i.e., you only care about what happens in the long run
- How will you (the row player) play in the following repeated games?
  - Hint: the other player will eventually best-respond to whatever you do

<table>
<thead>
<tr>
<th></th>
<th>4, 4</th>
<th>3, 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5, 3</td>
<td>0, 0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1, 0</th>
<th>3, 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2, 1</td>
<td>4, 0</td>
<td></td>
</tr>
</tbody>
</table>

- Note relationship to optimal strategies to commit to
- There is some work on learning strategies that are in equilibrium with each other [Brafman & Tennenholtz AIJ04]
Hawk-Dove Game

[Price and Smith, 1973]

<table>
<thead>
<tr>
<th></th>
<th>Dove</th>
<th>Hawk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dove</td>
<td>1, 1</td>
<td>0, 2</td>
</tr>
<tr>
<td>Hawk</td>
<td>2, 0</td>
<td>-1, -1</td>
</tr>
</tbody>
</table>

- Unique symmetric equilibrium: 50% Dove, 50% Hawk
Evolutionary game theory

- Given: a symmetric 2-player game

<table>
<thead>
<tr>
<th></th>
<th>Dove</th>
<th>Hawk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dove</td>
<td>1, 1</td>
<td>0, 2</td>
</tr>
<tr>
<td>Hawk</td>
<td>2, 0</td>
<td>-1, -1</td>
</tr>
</tbody>
</table>

- Population of players; players randomly matched to play game
- Each player plays a pure strategy
  \[ p_s = \text{fraction of players playing strategy } s \]
  \[ p = \text{vector of all fractions } p_s \text{ (the state)} \]
- Utility for playing \( s \) is \( u(s, p) = \sum_s p_s u(s, s') \)
- Players reproduce at rate proportional to their utility; their offspring play the same strategy
  \[ \frac{dp_s(t)}{dt} = p_s(t)(u(s, p(t)) - \sum_s p_s u(s', p(t))) \]
  - Replicator dynamic
- What are the steady states?
A steady state is stable if slightly perturbing the state will not cause us to move far away from the state.

Proposition: every stable steady state is a Nash equilibrium of the symmetric game.

Slightly stronger criterion: a state is asymptotically stable if it is stable, and after slightly perturbing this state, we will (in the limit) return to this state.
Evolutionarily stable strategies

[Price and Smith, 1973]

• Now suppose players play mixed strategies

• A (single) mixed strategy $\sigma$ is evolutionarily stable if the following is true:

  – Suppose all players play $\sigma$

  – Then, whenever a very small number of invaders enters that play a different strategy $\sigma'$,

    the players playing $\sigma$ must get strictly higher utility than those playing $\sigma'$ (i.e., $\sigma$ must be able to repel invaders)
Properties of ESS

- **Proposition.** A strategy $\sigma$ is evolutionarily stable if and only if the following conditions both hold:

  1. For all $\sigma'$, we have $u(\sigma, \sigma) \geq u(\sigma', \sigma)$ (i.e., symmetric Nash equilibrium)

  2. For all $\sigma' (\neq \sigma)$ with $u(\sigma, \sigma) = u(\sigma', \sigma)$, we have $u(\sigma, \sigma') > u(\sigma', \sigma')$

- **Theorem** [Taylor and Jonker 1978, Hofbauer et al. 1979, Zeeman 1980].

  Every ESS is asymptotically stable under the replicator dynamic. (Converse does not hold [van Damme 1987].)
Invasion (1/2)

- Given: population $P_1$ that plays $\sigma = 40\%$ Dove, 60$\%$ Hawk
- Tiny population $P_2$ that plays $\sigma' = 70\%$ Dove, 30$\%$ Hawk invades
- $u(\sigma, \sigma) = .16*1 + .24*2 + .36*(-1) = .28$ but $u(\sigma', \sigma) = .28*1 + .12*2 + .18*(-1) = .34$
- $\sigma'$ (initially) grows in the population; invasion is successful
Invasion (2/2)

\[
\begin{array}{c|cc}
 & \text{Dove} & \text{Hawk} \\
\hline
\text{Dove} & 1, 1 & 0, 2 \\
\text{Hawk} & 2, 0 & -1, -1
\end{array}
\]

- Now $P_1$ plays $\sigma = 50\%$ Dove, $50\%$ Hawk.
- Tiny population $P_2$ that plays $\sigma' = 70\%$ Dove, $30\%$ Hawk invades.
- $u(\sigma, \sigma) = u(\sigma', \sigma) = .5$, so second-order effect:
  - $u(\sigma, \sigma') = .35*1 + .35*2 + .15*(-1) = .9$ but $u(\sigma', \sigma') = .49*1 + .21*2 + .09*(-1) = .82$
- $\sigma'$ shrinks in the population; invasion is repelled.
### Rock-Paper-Scissors

<table>
<thead>
<tr>
<th></th>
<th>0, 0</th>
<th>-1, 1</th>
<th>1, -1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>1, -1</td>
<td>0, 0</td>
<td>-1, 1</td>
</tr>
<tr>
<td>-1, 1</td>
<td>1, -1</td>
<td>0, 0</td>
<td></td>
</tr>
<tr>
<td>1, -1</td>
<td>-1, 1</td>
<td></td>
<td>0, 0</td>
</tr>
</tbody>
</table>

- Only one Nash equilibrium (Uniform)
- \( u(\text{Uniform, Rock}) = u(\text{Rock, Rock}) \)
- No ESS
### “Safe-Left-Right”

<table>
<thead>
<tr>
<th></th>
<th>Safe</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safe</td>
<td>1, 1</td>
<td>1, 1</td>
<td>1, 1</td>
</tr>
<tr>
<td>Left</td>
<td>1, 1</td>
<td>0, 0</td>
<td>2, 2</td>
</tr>
<tr>
<td>Right</td>
<td>1, 1</td>
<td>2, 2</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

- Can 100% Safe be invaded?
- Is there an ESS?
**The ESS problem**

**Input:** symmetric 2-player normal-form game.

**Q:** Does it have an evolutionarily stable strategy?


---

**Thm.** ESS is **NP-hard**
[Etessami and Lochbihler 2004]

**Thm.** ESS is **Σ^p_2**-hard
[Etessami and Lochbihler 2004]

**Thm.** ESS is **Σ^p_2**
[Nisan 2006]

**Thm.** ESS is **Σ^p_2**-hard
[C. 2013]

---

**Thm.** ESS is **coNP-hard**
[Etessami and Lochbihler 2004]

---

**Thm.** ESS is in **Σ^p_2**
[Etessami and Lochbihler 2004]
The standard $\Sigma_2^P$-complete problem

**Input:** Boolean formula $f$ over variables $X_1$ and $X_2$

**Q:** Does there exist an assignment of values to $X_1$ such that for every assignment of values to $X_2$ $f$ is true?
Discussion of implications

• Many of the techniques for finding (optimal) Nash equilibria will not extend to ESS

• Evolutionary game theory gives a possible explanation of how equilibria are reached…
  … for this purpose it would be good if its solution concepts aren’t (very) hard to compute!
Learning in Stackelberg games

[Letchford, C., Munagala SAGT’09]

See also: Blum, Haghtalab, Procaccia [NIPS’14]

• Unknown follower payoffs

• Repeated play: commit to mixed strategy, see follower’s (myopic) response

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>1,?</td>
<td></td>
<td>3,?</td>
</tr>
<tr>
<td>D</td>
<td>2,?</td>
<td></td>
<td>4,?</td>
</tr>
</tbody>
</table>
Learning in Stackelberg games…

[Letchford, C., Munagala SAGT’09]

Theorem. Finding the optimal mixed strategy to commit to requires

\[ O(Fk \log(k) + dLk^2) \]

samples

- \( F \) depends on the size of the smallest region
- \( L \) depends on desired precision
- \( k \) is # of follower actions
- \( d \) is # of leader actions
Three main techniques in the learning algorithm

- Find one point in each region (using random sampling)
- Find a point on an unknown hyperplane
- Starting from a point on an unknown hyperplane, determine the hyperplane completely
Finding a point on an unknown hyperplane

Intermediate state

Step 1. Sample in the overlapping region

Step 2. Connect the new point to the point in the region that doesn’t match

Step 3. Binary search along this line

Region: R
Determining the hyperplane

Step 1. Sample a regular d-simplex centered at the point

Step 2. Connect d lines between points on opposing sides

Step 3. Binary search along these lines

Step 4. Determine hyperplane (and update the region estimates with this information)
In summary: CS/AI pushing at some of the boundaries of game theory.

- Learning in games
- Behavioral (humans playing games)
- Conceptual (e.g., equilibrium selection)
- Representation
- Computation
- CS work in game theory
- Game theory
Backup slides
Computational complexity theory

NP problems for which “yes” answers can be efficiently verified

P problems that can be efficiently solved (incl. linear programming [Khachiyan 1979])

NP-hard problems at least as hard as anything in NP

(This picture assumes $P \neq NP$.)

Matching pennies with a sensitive target

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L</strong></td>
<td>1, -1</td>
<td>-1, 1</td>
</tr>
<tr>
<td><strong>R</strong></td>
<td>-2, 2</td>
<td>1, -1</td>
</tr>
</tbody>
</table>

- If we play 50% L, 50% R, opponent will attack L
  - We get $0.5 \times 1 + 0.5 \times (-2) = -0.5$
- What if we play 55% L, 45% R?
- Opponent has choice between
  - L: gives them $0.55 \times (-1) + 0.45 \times 2 = 0.35$
  - R: gives them $0.55 \times 1 + 0.45 \times (-1) = 0.1$
- We get $-0.35 > -0.5$
Matching pennies with a sensitive target

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>1, -1</td>
<td>-1, 1</td>
</tr>
<tr>
<td>R</td>
<td>-2, 2</td>
<td>1, -1</td>
</tr>
</tbody>
</table>

- **Us**
- **Them**

- What if we play 60% L, 40% R?
- Opponent has choice between
  - L: gives them \(0.6 \times (-1) + 0.4 \times 2 = 0.2\)
  - R: gives them \(0.6 \times 1 + 0.4 \times (-1) = 0.2\)
- We get -.2 either way
- This is the **maximin** strategy
  - Maximizes our minimum utility
Let’s change roles

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>1, -1</td>
<td>-1, 1</td>
</tr>
<tr>
<td>R</td>
<td>-2, 2</td>
<td>1, -1</td>
</tr>
</tbody>
</table>

- Suppose **we** know **their** strategy
- If they play 50% L, 50% R,
  - We play L, we get \(0.5 \times 1 + 0.5 \times (-1) = 0\)
- If they play 40% L, 60% R,
  - If we play L, we get \(0.4 \times 1 + 0.6 \times (-1) = -0.2\)
  - If we play R, we get \(0.4 \times (-2) + 0.6 \times 1 = -0.2\)
- This is the **minimax** strategy

**von Neumann’s minimax theorem [1927]: maximin value = minimax value** (~LP duality)
Correlated equilibrium as Bayes-Nash equilibrium

<table>
<thead>
<tr>
<th></th>
<th>$\theta_2=1$</th>
<th>$\theta_2=2$</th>
<th>$\theta_2=3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1=1$</td>
<td>\begin{tabular}{</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>$\theta_1=2$</td>
<td>\begin{tabular}{</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>$\theta_1=3$</td>
<td>\begin{tabular}{</td>
<td>c</td>
<td>c</td>
</tr>
</tbody>
</table>
The Polynomial Hierarchy

$\exists^p L = \{ x \in \{0,1\}^* \mid (\exists w \in \{0,1\}^{\leq p(|x|)}) (x,w) \in L \}$

$\forall^p L = \{ x \in \{0,1\}^* \mid (\forall w \in \{0,1\}^{\leq p(|x|)}) (x,w) \in L \}$

$\exists^p C = \{ \exists^p L \mid p \text{ is a polynomial and } L \text{ in } C \}$

$\forall^p C = \{ \forall^p L \mid p \text{ is a polynomial and } L \text{ in } C \}$

$\Sigma_0^p = \Pi_0^p = P$

$\Sigma_{i+1}^p = \exists^p \Pi_i^p$

$\Pi_{i+1}^p = \forall^p \Sigma_i^p$

$\Delta_0^p = \Sigma_0^p = P = \Pi_0^p = \Delta_1^p$
The ESS-RESTRICTED-SUPPORT problem

**Input:** symmetric 2-player normal-form game, subset $T$ of the strategies $S$

**Q:** Does the game have an evolutionarily stable strategy whose support is restricted to (a subset of) $T$?
MINMAX-CLIQUE

proved $\Pi_2^P(=\text{co}\Sigma_2^P)$-complete by Ko and Lin [1995]

**Input:** graph $G = (V, E)$, sets $I$ and $J$, partition of $V$ into subsets $V_{ij}$ (for $i$ in $I$ and $j$ in $J$), number $k$

**Q:** Is it the case that for every function $t : I \rightarrow J$, $U_i V_{i,t(i)}$ has a clique of size $k$?

Thank you, compendium by Schaefer and Umans!
Illustration of reduction

<table>
<thead>
<tr>
<th>$s_{11}$</th>
<th>$s_{12}$</th>
<th>$s_{21}$</th>
<th>$s_{22}$</th>
<th>$s_{v_{11}}$</th>
<th>$s_{v_{12}}$</th>
<th>$s_{v_{21}}$</th>
<th>$s_{v_{22}}$</th>
<th>$s_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{11}$</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
</tr>
<tr>
<td>$s_{12}$</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
</tr>
<tr>
<td>$s_{21}$</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
</tr>
<tr>
<td>$s_{22}$</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
</tr>
<tr>
<td>$s_{v_{11}}$</td>
<td>3/2</td>
<td>0</td>
<td>3/2</td>
<td>3/2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$s_{v_{12}}$</td>
<td>0</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>$s_{v_{21}}$</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_{v_{22}}$</td>
<td>3/2</td>
<td>3/2</td>
<td>0</td>
<td>3/2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_0$</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>3/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Unrestricted support?

• Just duplicate all the strategies outside $T$…

• (Appendix: result still holds in games in which every pure strategy is the unique best response to some mixed strategy)
Theorem. Finding all of the hyperplanes necessary to compute the optimal mixed strategy to commit to requires $O(Fk \log(k) + dLk^2)$ samples

- $F$ depends on the size of the smallest region
- $L$ depends on desired precision
- $k$ is the number of follower actions
- $d$ is the number of leader actions