Extensive-Form Games

- Moves are done sequentially, not simultaneously
- Game forms a tree
- Nodes are labeled by players
- Leaves show payoffs
**Extensive vs. Normal Form**

Problem: Normal-form representation is exponential in the size of the extensive-form representation.

```
<table>
<thead>
<tr>
<th></th>
<th>L/L</th>
<th>L/R</th>
<th>R/L</th>
<th>R/R</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>2,4</td>
<td>2,4</td>
<td>5,3</td>
<td>5,3</td>
</tr>
<tr>
<td>R</td>
<td>1,0</td>
<td>3,2</td>
<td>1,0</td>
<td>3,2</td>
</tr>
</tbody>
</table>
```
**Extensive vs. Normal Form**

![Game Tree Diagram]

**Problem:** (ignore, nuclear war) is a Nash equilibrium, but threat isn’t credible!
Subgame-perfect equilibrium

- Each subtree forms a subgame
- A set of strategies is a subgame-perfect equilibrium if it is a Nash equilibrium in each subgame
- A player may be able to improve his equilibrium payoff by eliminating strategies!
DOOMSDAY MACHINE

https://youtu.be/2yfXgu37iyI
BACKWARD INDUCTION
**Backward Induction**

Subgame-perfect equilibrium!
**Example: Centipede Game**

Even subgame-perfect equilibrium can lead to strange outcomes!
CHECKERS IS SOLVED

• Zermelo’s Theorem [1913]: Either white can force a win, or black can force a win, or both sides can force a draw

• **Proof:** Backward induction

• Schaeffer solved the game in 2007, after 18 years of computation: It’s a tie!

• Checkers game tree has $10^{20}$ nodes; chess has $10^{40}$; go has $10^{170}$
In 2016, AlphaGo beat Lee Sedol, one of the strongest players in the history of go, in a 5-game match. A milestone that experts thought was a decade away. Combination of tree search techniques and deep reinforcement learning.
DIFFICULTY OF VARIOUS GAMES FOR COMPUTERS

EASY

SOLVED COMPUTERS CAN PLAY PERFECTLY

SOLVED FOR ALL POSSIBLE POSITIONS

TIC-TAC-TOE

NIM

GHOST (1989)

CONNECT FOUR (1995)

COSPAR

CHECKERS (2007)

SCRABBLE

COUNTERSINE

REVERS

GO

COMPUTERS CAN BEAT TOP HUMANS

COMPUTERS STILL loose to top humans

(BUT FOCUS ON RE: COULD CHANGE THIS)

CHESS

DARPA

STARCraft

GO

COMPUTERS MAY NEVER OUTPLAY HUMANS

MAO

SNAKES AND LADDER

SEVEN MINUTES IN HEAVEN

CALVINBALL
IMPERFECT-INFORMATION GAMES

• A chance node chooses between several actions according to a known probability distribution.

• An information set is a set of nodes that a player may be in, given the available information.

• A strategy must be identical for all nodes in an information set.
Example: Spaceship Game

- **Poll 1:** In Nash equilibrium, what is the expected payoff of player 1?
  1. 0.5
  2. 1
  3. 1.5
  4. 2
  5. 2.5
**Example: Spaceship Game**

\[ p = 0.5 \]

\[ p = 0.5 \]

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

C

Fold

Bet

H

T

-5,0

5,0

Fold

Bet

-1,1

1,-1

1,-1

-1,1
Example: Spaceship Game
Impossible to compute the optimal strategy of a subgame in isolation, unlike perfect info games!
INTERLUDE: ZERO-SUM GAMES
**INTERLUDE: ZERO-SUM GAMES**

- **Maximin** (randomized) strategy of player 1 maximizes the worst-case expected payoff
- In the penalty shot game, optimal strategy for both players is playing \( \left( \frac{1}{2}, \frac{1}{2} \right) \)
- In the game below, if shooter uses \((p, 1 - p)\):
  - Jump left: \(-\frac{p}{2} + 1 - p = 1 - \frac{3}{2}p\)
  - Jump right: \(p - 1 + p = 2p - 1\)
  - Maximize \(\min\{1 - \frac{3}{2}p, 2p - 1\}\) over \(p\)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-\frac{1}{2})</td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>
**INTERLUDE: ZERO-SUM GAMES**

- Denote the reward of player 1 from strategies \((s_1, s_2)\) by \(R(s_1, s_2)\)
- Maximin strategy is computed via LP:

\[
\begin{align*}
\text{max} & \quad w \\
\text{s.t.} & \quad \forall s_2 \in S, \sum_{s_1 \in S} p(s_1)R(s_1, s_2) \geq w \\
& \sum_{s_1 \in S} p(s_1) = 1 \\
& \forall s_1 \in S, p(s_1) \geq 0
\end{align*}
\]
INTERLUDE: THE MINIMAX THEOREM

• Theorem [von Neumann 1928]:
  Every 2-player zero-sum game has a unique value \( v \) such that:
  
  o Player 1 can guarantee value at least \( v \)
  o Player 2 can guarantee loss at most \( v \)

• Poll 2: How many Nash equilibrium payoffs do zero-sum games have?
  
  1. At most one
  2. At least one
  3. Exactly one

von Neumann
SOLVING IMPERFECT INFO GAMES

• Focus on zero-sum games (such as poker)
• We just saw that linear programming solves normal-form, zero-sum games in polynomial time
• But size of the normal-form game is exponential in the extensive-form representation!
• Work directly on extensive-form game
Solving Imperfect Info Games

- Player 1 constraints are linear:
  - $p_a + p_b = 1$
  - $p_c + p_d = 1$
  - $p_e + p_f = 1$
  - $\forall x, p_x \geq 0$

- Fix a strategy $q_\alpha, q_\beta$ for player 2, then the best response of player 1 is:
  \[
  \max_p 2p_b q_\alpha p_f - 2p_b q_\beta p_f - 2p_c q_\alpha + 6p_d
  \]
  which leads to a nonconvex problem!
SEQUENCE FORM

- **Insight:** last action taken by a player is the same for all nodes in an information set
  - **Perfect recall:** A player never forgets something he knew in the past
  - This is a restriction on the structure of the game
- Introduce scaled probability variables $p'_x$
- Information set constraint: $\sum_{x \in A_I} p'_x = p'_y$, where $A_I$ is the set of actions in information set $I$, and $y$ is the last action before reaching $I$
- To recover probabilities, set $p_x = p'_x / p'_y$
SEQUENCE FORM

- Player 1 constraints are linear:
  - \( p'_a + p'_b = 1 \)
  - \( p'_c + p'_d = 1 \)
  - \( p'_e + p'_f = p'_b \)
  - \( \forall x, p'_x \geq 0 \)
- Fix a strategy \( q_\alpha, q_\beta \) for player 2, then the best response of player 1 is:
  \[
  \max_p 2q_\alpha p'_f - 2q_\beta p'_f - 2p'_c q_\alpha + 6p'_d
  \]
  which is linear!
SEQUENCE FORM

- We showed how to compute a best response for a fixed opponent strategy
- **Fact:** Using “LP duality”, we can compute best responses for both players simultaneously
- **Fact:** This gives a method for computing optimal strategies
- Used to compute optimal strategies for Rhode Island Hold’em poker, which has roughly $10^8$ nodes [Gilpin and Sandholm 2007]
- But No Limit Texas Hold’em has $10^{167}$ nodes
April 24–May 8, 2015, at Rivers Casino, Pittsburgh
The first time a computer program has played human pros in a heads-up, no-limit poker game
SUMMARY

• Terminology:
  o Extensive-form game
  o Subgame perfect equilibrium
  o Imperfect information, information set
  o Perfect recall

• Algorithms:
  o Solving zero-sum games via LP
  o Sequence form-based approach to solving imperfect information games