• Assumptions so far:
  – Two-player game: Player A and B.
  – Perfect information: Both players see all the states and decisions. Each decision is made sequentially.
  – Zero-sum: Player’s A gain is exactly equal to player B’s loss.

• We are going to eliminate these constraints. We will eliminate first the assumption of “perfect information” leading to far more realistic models.
  – Some more game-theoretic definitions → Matrix games
  – Minimax results for perfect information games
  – Minimax results for hidden information games
A pure strategy for a player defines the move that the player would make for every possible state that the player would see.
Pure strategies for A:
Strategy I: (1→L,4→L)
Strategy II: (1→L,4→R)
Strategy III: (1→R,4→L)
Strategy IV: (1→R,4→R)

Pure strategies for B:
Strategy I: (2→L,3→L)
Strategy II: (2→L,3→R)
Strategy III: (2→R,3→L)
Strategy IV: (2→R,3→R)

In general: If $N$ states and $B$ moves, how many pure strategies exist?

Matrix form of games

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-1</td>
<td>-1</td>
<td>+2</td>
<td>+2</td>
</tr>
<tr>
<td>II</td>
<td>+4</td>
<td>+4</td>
<td>+2</td>
<td>+2</td>
</tr>
<tr>
<td>III</td>
<td>+5</td>
<td>+1</td>
<td>+5</td>
<td>+1</td>
</tr>
<tr>
<td>IV</td>
<td>+5</td>
<td>+1</td>
<td>+5</td>
<td>+1</td>
</tr>
</tbody>
</table>
• Matrix normal form of games: The table contains the payoffs for all the possible combinations of pure strategies for Player A and Player B.
• The table characterizes the game completely, there is no need for any additional information about rules, etc.
• Although, in many cases, the number of pure strategies may be too large for the table to be represented explicitly, the matrix representation is the basic representation that is used for deriving fundamental properties of games.
Minimax → Matrix version

For each strategy (each row of the game matrix), Player A should assume that Player B will use the optimal strategy given Player A’s strategy (the strategy with the minimum value in the row of the matrix). Therefore the best value that Player can achieve is the maximum over all the rows of the minimum values across each of the rows:

\[
\begin{array}{cccc}
\text{Max} & \text{Min} & M(i, j) \\
\text{Rows } i & \text{Columns } j
\end{array}
\]

The corresponding pure strategy is the optimal solution for this game → It is the optimal strategy for A assuming that B plays optimally.

Max value =

\[
\begin{array}{cccc}
\text{I} & \text{II} & \text{III} & \text{IV} \\
\text{I} & -1 & -1 & +2 & +2 \\
\text{II} & +4 & +4 & +2 & +2 \\
\text{III} & +5 & +1 & +5 & +1 \\
\text{IV} & +5 & +1 & +5 & +1
\end{array}
\]

\[
\text{Min value across each row}
\]

\[
\begin{array}{cccc}
\text{I} & \text{II} & \text{III} & \text{IV} \\
\text{Max value across each column} \\
\text{I} & -1 & -1 & +2 & +2 \\
\text{II} & +4 & +4 & +2 & +2 \\
\text{III} & +5 & +1 & +5 & +1 \\
\text{IV} & +5 & +1 & +5 & +1 \\
\text{Min of all the columns}
\end{array}
\]

\[
\begin{array}{cccc}
\text{Min} & \text{Max} & M(i, j) \\
\text{Columns } j & \text{Rows } i
\end{array}
\]

\[
\begin{array}{cccc}
\text{I} & \text{II} & \text{III} & \text{IV} \\
+5 & +4 & +5 & +2
\end{array}
\]
Minimax or Maximin?

- But we could have used the opposite argument:
- For each strategy (each column of the game matrix), Player B should assume that Player A will use the optimal strategy given Player B’s strategy (the strategy with the maximum value in the column of the matrix):

\[ \text{Min} \quad \text{Max} \quad M(i, j) \]

\[ \text{Columns } j \quad \text{Rows } i \]

- Therefore the best value that Player B can achieve is the minimum over all the columns of the maximum values across each of the columns
- Problem: Do we get to the same result??
- Is there always a solution?

Max value across each column

\[
\begin{array}{cccc}
   & I & II & III & IV \\
I & -1 & -1 & +2 & +2 \\
II & +4 & +4 & +2 & +2 \\
III & +5 & +1 & +5 & +1 \\
IV & +5 & +1 & +5 & +1 \\
\end{array}
\]

+5 +4 +5 +2

Min value =

game value = +2

Note that we find the same value and same strategies in both cases. Is that always the case?

Max value across each row

\[
\begin{array}{cccc}
   & I & II & III & IV \\
-1 & I & -1 & -1 & +2 & +2 \\
+2 & II & +4 & +4 & +2 & +2 \\
+1 & III & +5 & +1 & +5 & +1 \\
+1 & IV & +5 & +1 & +5 & +1 \\
\end{array}
\]

Min value across each row

\[
\begin{array}{cccc}
   & I & II & III & IV \\
+5 & I & -1 & -1 & +2 & +2 \\
+4 & II & +4 & +4 & +2 & +2 \\
+1 & III & +5 & +1 & +5 & +1 \\
+1 & IV & +5 & +1 & +5 & +1 \\
\end{array}
\]

Max value across each column

\[
\begin{array}{cccc}
   & I & II & III & IV \\
+5 & I & -1 & -1 & +2 & +2 \\
+4 & II & +4 & +4 & +2 & +2 \\
+1 & III & +5 & +1 & +5 & +1 \\
+1 & IV & +5 & +1 & +5 & +1 \\
\end{array}
\]

Min value =

game value = +2
Minimax vs. Maximin

• Fundamental Theorem I (von Neumann):
  – For a two-player, zero-sum game with perfect information:
    • There always exists an optimal pure strategy for each player
    • Minimax = Maximin

• Note: This is a game-theoretic formalization of the minimax search algorithm that we studied earlier.

Another (Seemingly Simple) Game

• The two Players A and B each have a coin
• They show each other their coin, choosing to show either head or tail.
• If they both choose head → Player B pays Player A $2
• If they both choose tail → Player B pays Player A $1
• If they choose different sides → Player A pays Player B $1
Side Note about all toy examples

• If you find this kind of toy example annoying, it models a large number of real-life situations.
• For example: Player A is a business owner (e.g., a restaurant, a plant..) and Player B is an inspector. The inspector picks a day to conduct the inspection; the owner picks a day to hide the bad stuff. Player B wins if the days are different; Player A wins if the days are the same.
• This class of problems can be reduced to the “coin game” (with different payoff distributions, perhaps).

Extensive Form
Problem: Since the moves are simultaneous, Player B does not know which move Player A chose. This is no longer a game with perfect information, so we have to deal with hidden information.

Extensive Form

Player A

Player B

+2
-1
-1
+1

Player B

Player A

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>+2</td>
<td>-1</td>
</tr>
<tr>
<td>T</td>
<td>-1</td>
<td>+1</td>
</tr>
</tbody>
</table>
Matrix Normal Form

- It is no longer the case that maximin = minimax (easy to verify: -1 vs. +1)
- Therefore: It appears that there is no pure strategy solution
- In fact, in general, *none of the pure strategies* are solutions to a zero-sum game with *hidden information*

Why no Pure Strategy Solutions?

- Intuition:
- If Player A considers move H, he has to assume that Player B will choose the worst-case move (for A), which is move T
  - Therefore Player A should try move T instead, but then he has to assume that Player B will choose the worst-case move (for A), which is move H.
  - Therefore Player A should consider move H, but then he has to assume that Player B will choose the worst-case move (for A), which is move T.
  - Therefore Player A should try move T instead, but then he has to assume that Player B will choose the worst-case move (for A), which is move H.
  - Therefore Player A should consider move H, but then he has to assume that Player B will choose the worst-case move (for A), which is move T.
Using Random Strategies

- Suppose that, instead of choosing a fixed pure strategy, Player A chooses randomly strategy $H$ with probability $p$, and strategy $T$ with probability $1-p$.
- If Player B chooses move $H$, the expected payoff for Player A is:
  $$p \times (+2) + (1-p) \times (-1) = 3p - 1$$
- If Player B chooses move $T$, the expected payoff for Player A is:
  $$p \times (-1) + (1-p) \times (+1) = -2p + 1$$
- So, the worst case is when Player B chooses a strategy that minimizes the payoff between the 2 cases:
  $$\min(3p - 1, -2p + 1)$$
- Player A should then adjust the probability $p$ so that its payoff is maximized (note the similarity with the standard maximin procedure described earlier):
  $$\max_p \min(3p - 1, -2p + 1)$$

Graphical Solution

- Expected payoff if Player B chooses $T$: $-2p + 1$
- Expected payoff if Player B chooses $H$: $3p - 1$
Graphical Solution

- Expected payoff if Player B chooses T:
  \[-2p + 1\]

- Expected payoff if Player B chooses H:
  \[3p - 1\]

No matter what strategy Player B follows (choosing a move at random with prob. \(q\) for H), the resulting payoffs will still be between the two lines corresponding to B’s pure strategies.

- Excepted payoff for Player A:
  \[-1\]
Mixed Strategies

- It is no longer possible to find an optimal pure strategy for Player A.
- We need to change the problem a bit: We assume that Player A chooses a pure strategy randomly at the beginning of the game.
- In that scenario, Player A selects one pure strategy probability \( p \) and the other one with probability \( 1-p \).
- This strategy of choosing pure strategies randomly is called a mixed strategy for Player A and is entirely defined by the probability \( p \).
- Question: We know that we cannot find an optimal pure strategy for Player A, but can we find an optimal mixed strategy \( p \)?
- Answer: Yes! The result that we derived for the simple example holds for general games. It yields a procedure for finding the optimal mixed strategy for zero-sum games.
Minimax with Mixed Strategies

• Theorem II (von Neumann):
  – For a two-player, zero-sum game with hidden information:
    • There always exists an optimal mixed strategy with value
    \[
    \max p \min (p \times m_{11} + (1 - p) \times m_{21}, p \times m_{12} + (1 - p) \times m_{22})
    \]
    • Where the matrix form of the game is:
      \[
      \begin{bmatrix}
      m_{11} & m_{12} \\
      m_{21} & m_{22}
      \end{bmatrix}
      \]
    • Note: This is a direct generalization of the minimax result to mixed strategies.

Minimax with Mixed Strategies

• Theorem II (von Neumann):
  – For a two-player, zero-sum game with hidden information:
    • There always exists an optimal mixed strategy
    • In addition, just like for games with perfect information, it does not matter in which order we look at the players, minimax is the same as maximin
    \[
    \max p \min (p \times m_{11} + (1 - p) \times m_{21}, p \times m_{12} + (1 - p) \times m_{22}) = \\
    \min q \max (q \times m_{11} + (1 - q) \times m_{21}, q \times m_{12} + (1 - q) \times m_{22})
    \]
    • Note: This is a direct generalization of the minimax result to mixed strategies.
Recipe for 2x2 games

\[
\min(p \times m_{11} + (1 - p) \times m_{21}, p \times m_{12} + (1 - p) \times m_{22})
\]

- Since the two functions of \( p \) are linear, the maximum is attained either for:
  - \( p = 0 \)
  - \( p = 1 \)
  - The intersection of the two lines, if it occurs for \( p \) between 0 and 1
General Case: \( N \times M \) Games

- We have illustrated the problem on 2x2 games (2 strategies for each of Player A and Player B)
- The result generalizes to \( N \times M \) games, although it is more difficult to compute
- A mixed strategy is a vector of probabilities (summing to 1!) \( p = (p_1, ..., p_N) \). \( p_i \) is the probability with which strategy \( i \) will be chosen by Player A.
- The optimal strategy is found by solving the problem:

\[
\max_p \min_j \sum_i p_i \cdot m_{ij}
\]

\[
\sum_i p_i = 1
\]

This is solved by using “Linear Programming.”

Expected payoff for Player A if Player B chooses pure strategy number \( j \) and Player A chooses pure strategy \( i \) with prob. \( p_i \).
Discussion

- The criterion for selecting the optimal mixed strategy is the average payoff that Player A would receive over many runs of the game.
- It may seem strange to use random choices of pure strategies as “mixed” strategies and to search for optimal mixed strategies.
- In fact, it formalizes what happens in common situations. For example, in poker, if Player A follows a single pure strategy (taking the same action every time a particular configuration of cards is dealt), Player B can guess and respond to that strategy and lower Player A’s payoffs. The right thing to do is for Player A to change randomly the way each configuration is handled, according to some policy. A good player would use a good policy.
- The game theory results formalize the need for things like “bluffing” in games with hidden information.
- The theory assumes rational players → Roughly speaking, the players make decisions based on increasing their respective payoffs (utility values, preferences,...).
Another Example: Sort of Poker

- Players A and B play with two types of cards: Red and Black
- Player A is dealt one card at random (50% prob. of being Red)
- If the card is red, Player A may resign and loses $10
- Or Player A may hold
  - B may then resign → A wins $4
  - B may see
    - A loses $20 if the card is Red
    - A wins $16 otherwise

Modified version of an example from Andrew Moore
The game is non-deterministic because of the initial random choice of cards.

Hidden information: Player B cannot know which of these 2 states it’s in.

• Generate the matrix form of the game (be careful: It’s not a deterministic game)
• Find the optimal mixed strategy
• Find the expected payoff for Player A
Summary

• Matrix form of games
• Minimax procedure and theorem for games with perfect information \( \Rightarrow \) Always a \textit{pure strategy} solution
• Minimax procedure and theorem for games with perfect information \( \Rightarrow \) Always a \textit{mixed strategy} solution
• Procedure for solving 2x2 games with hidden information
• Understanding of how the problem is formalized for \( NxM \) games (actually solving them requires linear programming tools which will not be covered here)

• Important: These results apply only to \textit{zero-sum games}. This is still a severe restrictions as most realistic decision-making problems cannot be modeled as \textit{zero-sum games} \( \Rightarrow \) This restriction will be eliminated next!