CHAPTER 11: MULTIAgENT INTERACTIONS

An Introduction to Multiagent Systems

http://www.csc.liv.ac.uk/~mjw/pubs/imas/

1. What are Multiagent Systems?

2. Utilities and Preferences

Thus a multiagent system contains a number of agents...
Utility functions lead to preference orderings over outcomes:

- $\omega \succeq \omega'$ means $u_i(\omega) \geq u_i(\omega')$
- $\omega \succ \omega'$ means $u_i(\omega) > u_i(\omega')$

Utility functions lead to preference orderings over outcomes:

- $(\mathcal{M})!n < (\mathcal{M})!n$ means $\mathcal{M} ! \prec \mathcal{M}$
- $(\mathcal{M})!n \preceq (\mathcal{M})!n$ means $\mathcal{M} ! \preceq \mathcal{M}$
Environment behaviour given by state transformer function:

\[
\tau: \text{A} \times \text{A} \to \Omega
\]

- Here is a state transformer function:
  \[
  \tau(\text{D}, \text{D}) = \omega_1 \quad \tau(\text{D}, \text{C}) = \omega_2 \quad \tau(\text{C}, \text{D}) = \omega_3 \quad \tau(\text{C}, \text{C}) = \omega_4
  \]
  (This environment is sensitive to actions of both agents.)

- Here is another:
  \[
  \tau(\text{D}, \text{D}) = \omega_1 \quad \tau(\text{D}, \text{C}) = \omega_1 \quad \tau(\text{C}, \text{D}) = \omega_1 \quad \tau(\text{C}, \text{C}) = \omega_1
  \]
  (Neither agent has any influence in this environment.)

- And here is another:
  \[
  \tau(\text{D}, \text{D}) = \omega_1 \quad \tau(\text{D}, \text{C}) = \omega_2 \quad \tau(\text{C}, \text{D}) = \omega_1 \quad \tau(\text{C}, \text{C}) = \omega_2
  \]
  (Neither agent has any influence in this environment.)

- Here is another:
  \[
  \tau(\text{D}, \text{D}) = \omega_1 \quad \tau(\text{D}, \text{C}) = \omega_2 \quad \tau(\text{C}, \text{D}) = \omega_2 \quad \tau(\text{C}, \text{C}) = \omega_2
  \]
  (This environment is controlled by \text{j}.)

Rational Action

- Suppose we have the case where both agents can influence the outcome, and they have utility functions as follows:
  \[
  u_i(\omega_1) = 1 \quad u_i(\omega_2) = 1 \quad u_i(\omega_3) = 4 \quad u_i(\omega_4) = 4
  \]
  \[
  u_j(\omega_1) = 1 \quad u_j(\omega_2) = 4 \quad u_j(\omega_3) = 1 \quad u_j(\omega_4) = 4
  \]

- With a bit of abuse of notation:
  \[
  u_i(\text{D}, \text{D}) = 1 \quad u_i(\text{D}, \text{C}) = 1 \quad u_i(\text{C}, \text{D}) = 4 \quad u_i(\text{C}, \text{C}) = 4
  \]

- Then agent \text{i}'s preferences are:
  \[
  \text{C} \succ \text{C}, \text{C} \succ \text{D}, \text{D} \succ \text{C}, \text{D} \succ \text{D}, \text{D}
  \]
  \text{“C” is the rational choice for \text{i}.} (Because \text{i} prefers all outcomes that arise through \text{C} overall outcomes that arise through \text{D}).
Payoff Matrices

- We can characterise the previous scenario in a payoff matrix.
- In Nash equilibrium, there isn't always a dominant strategy:
  - We will say that a strategy $s_i$ is dominant for player $i$ if at least as well playing $s_i$ as it would doing anything else.
  - Agent $i$ is the row player.
  - Agent $j$ is the column player.

Dominant Strategies

- In Nash equilibrium, neither agent has any incentive to deviate from a
  - Agent $j$ is the row player.
  - Agent $j$ is the column player.

Solution Concepts

- How will a rational agent will behave in any given scenario?
- Answered in solution concepts:
  - Nash equilibrium strategies;
  - $\begin{array}{c|c}
    \text{coop} & 4 \\
    \hline
    \text{defect} & 1 \\
  \end{array}$
  - $\begin{array}{c|c}
    \text{coop} & 1 \\
    \hline
    \text{defect} & 4 \\
  \end{array}$
  - $\begin{array}{c|c}
    \text{coop} & 4 \\
    \hline
    \text{defect} & 1 \\
  \end{array}$

- Strategies that maximise social welfare:
  - Pareto optimal strategies;
  - Nash equilibrium strategies;
  - Dominant strategies:

- Unfortunately, there isn't always a dominant strategy:
  - Agent $i$ is the row player.
  - Agent $j$ is the column player.

Solution Concepts

- We can characterise the previous scenario in a payoff matrix.
  - Agent $i$ is the row player.
  - Agent $j$ is the column player.

Dominant Strategies

- In Nash equilibrium, neither agent has any incentive to deviate from a
  - Agent $i$ is the row player.
  - Agent $j$ is the column player.

Solution Concepts

- How will a rational agent will behave in any given scenario?
- Answered in solution concepts:
  - Nash equilibrium strategies;
  - $\begin{array}{c|c}
    \text{coop} & 4 \\
    \hline
    \text{defect} & 1 \\
  \end{array}$
  - $\begin{array}{c|c}
    \text{coop} & 1 \\
    \hline
    \text{defect} & 4 \\
  \end{array}$
  - $\begin{array}{c|c}
    \text{coop} & 4 \\
    \hline
    \text{defect} & 1 \\
  \end{array}$

- Strategies that maximise social welfare:
  - Pareto optimal strategies;
  - Nash equilibrium strategies;
  - Dominant strategies:

- Unfortunately, there isn't always a dominant strategy:
  - Agent $i$ is the row player.
  - Agent $j$ is the column player.
1. Not every interaction scenario has a Nash equilibrium.
2. Some interaction scenarios have more than one equilibrium.

Matching Pennies: The Payoff Matrix

<table>
<thead>
<tr>
<th></th>
<th>Heads</th>
<th>Tails</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heads</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Tails</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>
Chapter 11: An Introduction to Multiagent Systems

2e

**Mixed Strategies**

A mixed strategy has the form:

- play $\alpha_1$ with probability $p_1$
- play $\alpha_2$ with probability $p_2$
- ...
- play $\alpha_k$ with probability $p_k$

such that $p_1 + p_2 + ... + p_k = 1$.

Nash proved that every finite game has a Nash equilibrium in mixed strategies.

So this result overcomes the lack of solutions for pure strategies (unlike the case for Nash equilibrium in mixed strategies).

If an outcome is not Pareto optimal, then there is another outcome that makes everyone better off (i.e. no other outcome makes one agent better off without making another agent worse off).

Nash's Theorem

If an outcome is not Pareto optimal, then there is another outcome that makes everyone better off (i.e. no other outcome makes one agent better off without making another agent worse off).

Pareto Optimality

If an outcome is not Pareto optimal, then there is another outcome that makes everyone better off (i.e. no other outcome makes one agent better off without making another agent worse off).

A mixed strategy has the form:

- play $\alpha_1$ with probability $p_1$
- ...
- play $\alpha_k$ with probability $p_k$
Chapter 11: An Introduction to Multiagent Systems

2. Social Welfare

Social Welfare

• The social welfare of an outcome \( \omega \) is the sum of the utilities that each agent gets from \( \omega \):

\[
\sum_{i \in \mathcal{A}} u_i(\omega)
\]

• Think of it as the "total amount of money in the system".

• As a solution concept, may be appropriate when the agent does not have the same preferences and zero-sum interactions are those where utilities sum to zero: 

\[ u_i(\omega) = (\pi_i) + (\pi_j) \]

2. Competitive and Zero-Sum Interactions

• Where preferences of agents are diametrically opposed we have strictly competitive scenarios. Zero-sum encounters are those where utilities sum to zero: competitive and zero-sum interactions.

\[ (\pi_i) + (\pi_j) \]

2. The Prisoner’s Dilemma

Two men are collectively charged with a crime and held in separate cells, with no way of meeting or communicating.

• People frequently act as if they were in a zero-sum game, even if the individual is not.

• Zero-sum encounters in real life are very rare... but

\[ u_i(\omega) + u_j(\omega) = 0 \]

for all \( \omega \in \Omega \).

4. The Prisoner’s Dilemma

Two prisoners know that if neither confesses, both will each be jailed for one year.

• If both confess, then each will be jailed for two years.

• If one confesses and the other does not, the one who confesses will be free, and the other will be jailed for three years.

They are told that:

The outcome is the worst for you unless you have got negative utility. The best for you is the worst for the other.

For all \( \omega \in \Omega \):

\[ 0 = (\pi_i) + (\pi_j) \]

As a solution concept, may be appropriate when the whole system (all agents) has a single owner (then the system is important, not individuals).

Think of it as the "total amount of money in the system".

The social welfare of an outcome \( \omega \) is the sum of the utilities that each agent gets from \( \omega \):

\[ \text{Social Welfare} \]
• Payoff matrix for prisoner’s dilemma:

<table>
<thead>
<tr>
<th></th>
<th>coop</th>
<th>defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>coop</td>
<td>3, 3</td>
<td>0, 5</td>
</tr>
<tr>
<td>defect</td>
<td>5, 0</td>
<td>2, 2</td>
</tr>
</tbody>
</table>

- Topleft: If both defect, then both get punishment for mutual defection.
- Top right: If both cooperate, then mutual cooperation.
- Bottom right: Reward for mutual cooperation.
- Bottom left: If both cooperate and defect, defect gets a sucker’s payoff of 1, while coop gets 4.

What Should You Do?

• The individual rational action is defect. Surely they should both cooperate and each get payoff of 3! But intuition says this is not the best outcome: both agents defect, and get payoff = 2.

Solution Concepts

• D is a dominant strategy.
• (D, D) is the only Nash equilibrium.
• All outcomes except (D, D) are Pareto optimal.

• (C, C) maximises social welfare.
• So defection is the best response to all possible strategies: both agents defect, and get payoff = 2. This guarantees a payoff of no worse than 2, whereas cooperating guarantees a payoff of at most 1.

• (C, C) maximises social welfare.
• (D, D) is a dominant strategy.
Chapter 11: An Introduction to Multiagent Systems

This apparent paradox is the fundamental problem of multi-agent interactions.

Real-world examples:
- Nuclear arms reduction ("why don't I keep mine..."
- Free rider systems — public transport.

The prisoner's dilemma is ubiquitous.

Can we recover cooperation?

- The shadow of the future...
- Program equilibria and mediators
- The prisoner's dilemma is ambiguous.
- The other prisoner is my twin!
- We are not all madmen!!!

Arguments to recover cooperation:
- Somehow the dilemma is being formulated wrongly.
- The game theory notion of rational action is wrong.
- Conclusions that some have drawn from this analysis.

Arguments for Reconciling Cooperation

Consider the following program:

4.1 Program Equilibria

The best response to this program is to submit the program.

Here == is lexical comparison.

END IF.
DO(D)!
ELSE
DO(C)!

IF H1's program == H2's program
THEN

4.2 Program Equilibria

The strategic notion of rational action is misleading.

■ Conclusions that some have drawn from this analysis.

■ Arguments for Reconciling Cooperation

End-IF.

Consider the following program:

IF H1's program == This program
THEN
DO(C);
ELSE
DO(D);
END-IF.

Here == is textual comparison.

The best response to this program is to submit the same program, giving an outcome of (C, C).
4.3 The Iterated Prisoner's Dilemma

One answer: play the game more than once.

But... suppose you both know that you will play the game exactly \( n \) times. On round \( n - 1 \), you have an incentive to defect, too. But this makes round \( n - 2 \) the last “real”, and so you have an incentive to defect, too. This is the backwards induction problem.

Playing the prisoner's dilemma with a fixed, finite, pre-determined, commonly known number of rounds, defection is the best strategy.

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Axelrod (1984) investigated this problem, with a range of opponents... What strategy should you choose, so as to maximise your overall payoff?

Suppose you play Iterated Prisoner's Dilemma against a range of opponents... What strategy should you choose, so as to maximise your overall payoff?
Chapter 11
An Introduction to Multiagent Systems

2e

Strategies in Axelrod’s Tournament

• ALLD: “Always defect”—the hawk strategy;

• TIT-FOR-TAT:
  1. On round \( u = 0 \), cooperate.
  2. On round \( u > 0 \), do what your opponent did on round \( u - 1 \).

• TESTER:
  1. On round \( n = 0 \), cooperate.
  2. On round \( n < 0 \), do what your opponent did on round \( n \).
  3. On round \( n = 0 \), defect; if the opponent retaliated, then:

• JOSS:
  As TIT-FOR-TAT, except periodically defect.

Strategies in Axelrod’s Tournament

Recipes for Success in Axelrod’s Tournament

Axelrod suggests the following rules for succeeding in his tournament:

• Don’t be envious: Don’t play as if it were zero sum!

• Be nice: Don’t play as if it were zero sum!

• Retaliate appropriately:

• Don’t hold grudges: Always reciprocate cooperation immediately.

• Always reward cooperation immediately.

• Don’t be envious.

http://www.csc.liv.ac.uk/~mjw/pubs/imas/
Consider another type of encounter—the game of chicken:

<table>
<thead>
<tr>
<th></th>
<th>coop</th>
<th>defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>coop</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>defect</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

(Think of James Dean in Rebel without a Cause: swerving = coop, driving straight = defect.)

**Chicken:**
- mutual defection is most feared outcome.
- whereas sucker’s payoff is most feared in prisoners dilemma.

**Prisoner’s Dilemma:**
- All outcomes except (a’, a’) (where a’, a’ are the payoffs) are Pareto optimal.
- All outcomes except (a’, a’) are Nash equilibriums.
- Strategic pairs (a’, a’) and (a, a’) are Nash.
- There is no dominant strategy (in our sense).

**Solution Concepts**

6 Other Symmetric 2x2 Games:

- cooperation dominates games, there are 24 possible orderings on outcomes.

- Given the 4 possible outcomes of (symmetric) cooperate/defect games, there are 24 possible outcomes.

- There is no dominant strategy (in our sense).

- There are 24 possible orderings on outcomes.

- All outcomes except (D, D) are Pareto optimal.

- All outcomes except (D, D) are Nash equilibriums.

- There is no dominant strategy (in our sense).
\( CC \succ_i DC \succ_i DD \succ_i CD \)

*Stag hunt.*