Game Theory: Strategic decision making

29 March, 2017

Janyl Jumadinova
Multi-agent Systems
Multi-agent Interactions

- We are going to look at interactions in which agents affect each other through their actions.
- Assume agents to have “spheres of influence” that they control in the environment.
- Also, we assume that the welfare (goal achievement, utility) of each agent at least partially depends on the actions of others.
- **What agents should do in the presence of other agents (which also do stuff)?**
Utilities and Preferences

- Assume $O = \{o_1; ..., o_n\}$ is a set of possible outcomes (e.g. possible “runs” of the system until final states are reached).
- Agents are assumed to be self-interested: they have preferences over how the environment is.
Utilities and Preferences

- Assume $O = \{o_1; \ldots, o_n\}$ is a set of possible outcomes (e.g. possible “runs” of the system until final states are reached).
- Agents are assumed to be self-interested: they have preferences over how the environment is.
- We capture preferences by utility functions:
  
  $u_i : O \rightarrow R$
  $u_j : O \rightarrow R$
Utilities and Preferences

- Assume $O = \{o_1; \ldots, o_n\}$ is a set of possible outcomes (e.g., possible “runs” of the system until final states are reached).

- Agents are assumed to be self-interested: they have preferences over how the environment is.

- We capture preferences by utility functions:
  \[ u_i : O \rightarrow R \]
  \[ u_j : O \rightarrow R \]

- Utility functions lead to preference orderings over outcomes:
  \[ o \succeq_i o' \text{ means } u_i(o) \geq u_i(o') \]
Utility

- Utility is not money (but it is a useful analogy)
- Typical relationship between utility and money:
Multi-agent Encounters

- We need a model of the environment in which these agents will act
  - agents simultaneously choose an action to perform, and as a result of the actions they select, an outcome in $O$ will result
  - the actual outcome depends on the combination of actions assume each agent has just two possible actions that it can perform, C and D
We need a model of the environment in which these agents will act

- agents simultaneously choose an action to perform, and as a result of the actions they select, an outcome in $O$ will result
- the actual outcome depends on the combination of actions assume each agent has just two possible actions that it can perform, C and D

Environment behavior given by state transformer function:
$$\tau : Ac \times Ac \rightarrow O$$
Here is a state transformer function:

\[ \tau(D, D) = o_1, \quad \tau(D, C) = o_2, \quad \tau(C, D) = o_3, \quad \tau(C, C) = o_4 \]

- This environment is sensitive to actions of both agents.
Here is a state transformer function:
\[ \tau(D, D) = o_1, \tau(D, C) = o_2, \tau(C, D) = o_3, \tau(C, C) = o_4 \]
- This environment is sensitive to actions of both agents.

Here is another:
\[ \tau(D, D) = o_1, \tau(D, C) = o_1, \tau(C, D) = o_1, \tau(C, C) = o_1 \]
- Neither agent has any influence in this environment.
Here is a state transformer function:
\[ \tau(D, D) = o_1, \tau(D, C) = o_2, \tau(C, D) = o_3, \tau(C, C) = o_4 \]
- This environment is sensitive to actions of both agents.

Here is another:
\[ \tau(D, D) = o_1, \tau(D, C) = o_1, \tau(C, D) = o_1, \tau(C, C) = o_1 \]
- Neither agent has any influence in this environment.

And here is another:
\[ \tau(D, D) = o_1, \tau(D, C) = o_2, \tau(C, D) = o_1, \tau(C, C) = o_2 \]
Here is a state transformer function:
\[ \tau(D, D) = o_1, \tau(D, C) = o_2, \tau(C, D) = o_3, \tau(C, C) = o_4 \]
- This environment is sensitive to actions of both agents.

Here is another:
\[ \tau(D, D) = o_1, \tau(D, C) = o_1, \tau(C, D) = o_1, \tau(C, C) = o_1 \]
- Neither agent has any influence in this environment.

And here is another:
\[ \tau(D, D) = o_1, \tau(D, C) = o_2, \tau(C, D) = o_1, \tau(C, C) = o_2 \]
- This environment is controlled by agent \( j \).
Suppose we have the case where both agents can influence the outcome, and they have utility functions as follows:

\[ u_i(o_1) = 1, u_i(o_2) = 1, u_i(o_3) = 4, u_i(o_4) = 4 \]
\[ u_j(o_1) = 1, u_j(o_2) = 4, u_j(o_3) = 1, u_j(o_4) = 4 \]

With a bit of abuse of notation:

\[ u_i(D, D) = 1, u_i(D, C) = 1, u_i(C, D) = 4, u_i(C, C) = 4 \]
\[ u_j(D, D) = 1, u_j(D, C) = 4, u_j(C, D) = 1, u_j(C, C) = 4 \]
Suppose we have the case where both agents can influence the outcome, and they have utility functions as follows:

\[ u_i(o_1) = 1, \quad u_i(o_2) = 1, \quad u_i(o_3) = 4, \quad u_i(o_4) = 4 \]

\[ u_j(o_1) = 1, \quad u_j(o_2) = 4, \quad u_j(o_3) = 1, \quad u_j(o_4) = 4 \]

With a bit of abuse of notation:

\[ u_i(D, D) = 1, \quad u_i(D, C) = 1, \quad u_i(C, D) = 4, \quad u_i(C, C) = 4 \]

\[ u_j(D, D) = 1, \quad u_j(D, C) = 4, \quad u_j(C, D) = 1, \quad u_j(C, C) = 4 \]

Then agent \( i \)'s preferences are: \( C, C \succ_i C, D \succ_i D, C \succ_i D, D \)
Suppose we have the case where both agents can influence the outcome, and they have utility functions as follows:

\[ u_i(o_1) = 1, \quad u_i(o_2) = 1, \quad u_i(o_3) = 4, \quad u_i(o_4) = 4 \]
\[ u_j(o_1) = 1, \quad u_j(o_2) = 4, \quad u_j(o_3) = 1, \quad u_j(o_4) = 4 \]

With a bit of abuse of notation:

\[ u_i(D, D) = 1, \quad u_i(D, C) = 1, \quad u_i(C, D) = 4, \quad u_i(C, C) = 4 \]
\[ u_j(D, D) = 1, \quad u_j(D, C) = 4, \quad u_j(C, D) = 1, \quad u_j(C, C) = 4 \]

Then agent \( i \)’s preferences are:

\[ C, C \succeq_i C, D \not\succeq_i D, C \succeq_i D, C \]

“\( C \)” is a **rational** choice for agent \( i \)

- Because \( i \) prefers all outcomes that arise through \( C \) over all outcomes that arise through \( D \).
We can characterize the previous scenario in a payoff matrix:

<table>
<thead>
<tr>
<th></th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>defect</td>
</tr>
<tr>
<td>defect</td>
<td>1</td>
</tr>
<tr>
<td>j</td>
<td>1</td>
</tr>
<tr>
<td>coop</td>
<td>1</td>
</tr>
</tbody>
</table>

Agent $i$ is the column player
Agent $j$ is the row player
Game Theory: Solution Concepts

How will a rational agent behave in any given scenario?

- Strategic decision making.
- Look for equilibrium (Nash in non-cooperative games), dominant, Pareto optimal, social welfare maximizing strategies (to discuss next time).
- In cases with more than one equilibrium: mixed strategies.
  - Choose one strategy with some probability
Inductive Reasoning

- A lot of the problems are not always ‘normal’, stationary nor in equilibrium.
Inductive Reasoning

- A lot of the problems are not always ‘normal’, stationary nor in equilibrium.
- Humans use inductive reasoning in complicated or ill-defined situations:
  - look for patterns,
  - make deductions based on current hypotheses and act on them,
  - modify our beliefs based on the environment’s feedback.
Inductive Reasoning

▶ A lot of the problems are not always ‘normal’, stationary nor in equilibrium.
▶ Humans use inductive reasoning in complicated or ill-defined situations:
  ▶ look for patterns,
  ▶ make deductions based on current hypotheses and act on them,
  ▶ modify our beliefs based on the environment’s feedback.
▶ We use models to fill gaps in our understanding.
A lot of the problems are not always ‘normal’, stationary nor in equilibrium.

Humans use inductive reasoning in complicated or ill-defined situations:

- look for patterns,
- make deductions based on current hypotheses and act on them,
- modify our beliefs based on the environment’s feedback.

We use models to fill gaps in our understanding.

We can apply inductive reasoning in software agents as well!
Many real-life situations involve a set of independent agents/entities competing for the same resource, in an uncoordinated fashion:

- drivers choosing similar travel routes,
- visitors to a popular website,
- ...,
- wireless devices (wifi, Bluetooth etc) sharing RF spectrum.
$N$ people decide independently each week whether to go to a bar that offers entertainment on a certain night.

- Assume $N$ is set at 100.
- Space is limited, and the evening is enjoyable if things are not too crowded. Especially, if fewer than 60 percent of the possible 100 are present.
- There is no sure way to tell the numbers coming in advance; therefore a person or an agent goes (deems it worth going) if he expects fewer than 60 to show up or stays home if he expects more than 60 to go.
El Farol bar game

- Each customer has a finite set of predictors (strategies) which s/he uses to predict next week’s attendance.
- Each predictor has a score associated with it.
- Customers use the best predictor to predict next week’s attendance and to decide to stay or go.
El Farol bar game

...44 78 56 15 23 67 84 34 45 76 40 56 22 35

The same as last week  35 ⇒ go

A (rounded) average of the last m attendances.  49 ⇒ go

The same as 3 weeks ago.  56 ⇒ go

The trend in the last 8 weeks (bounded by 0 and 100)  76 ⇒ stay

...
El Farol bar game: Implementation

- Agents determine if they will attend the bar based on their strategies and the prediction of the attendance for the current time step.
El Farol bar game: Implementation

- Agents determine if they will attend the bar based on their strategies and the prediction of the attendance for the current time step.
- Each strategy is given a score based on the knowledge of the agent (i.e., based on the current state recorded in the memory of the agent). The smaller value returned by this function the better the strategy.
- The score value is constructed by summing up the differences between the actual and predicted attendance values during the past ‘memorySize’ weeks.
El Farol bar game: Implementation

- The predicted value of the attendance is determined by:
  \[ p(t) = w(t) + \sum_{i=t-1}^{t-M} w(i) \times a(i - 1), \]
  where \( w \) is the weight and \( a \) is the attendance level.
El Farol bar game: Implementation

- The predicted value of the attendance is determined by:
  \[ p(t) = w(t) + \sum_{i=t-1}^{t-M} w(i) \times a(i-1), \]
  where \( w \) is the weight and \( a \) is the attendance level.

- Weights are generated randomly (-1, 1) and make up the strategy.

- Then the global history is updated.

- At the end of the time step, when new attendance level is known, all agents update their best strategy based on the scores.