A Language for Opponent Modeling in Repeated Games

Ya’akov Gal and Avi Pfeffer
Harvard University

Opponent Modeling

• Agent modeler describes patterns in behavior of opponents.
• Modeler plays a best-response to her model.
• Modeler does not assume that players are rational.
  – Players might be using heuristics, e.g. divide and conquer.
  – Players are uncertain about game structure.

Game theory and multi-agent systems

• Game theory has had a deep impact on AI.
• Provides a language for decision-making under uncertainty.
• Given game description, agents’ optimal decisions can be derived rationally (Nash equilibrium)

The problem with classical game theory

• Strong assumptions
  – Common knowledge of rationality, strategies and payoffs.
  – Game structure is known a priory.
• Unnatural representation of patterns and heuristics.
• Does not distinguish real-world vs. mental models used for deliberation.
Contributions

• A natural language for opponent modeling in games.
• Distinguishes between the real-world and possible worlds used for deliberation.
• Framework for learning opponents’ strategies (even non-stationary).

Networks of Influence Diagrams (NIDs)

• Graphical language where each node is a Multi Agent Influence Diagram (MAID) (Koller and Milch, 2001)
• MAIDs
  – Graphically describe a decision process where all agents know a correct model of the world.
  – Compute a Nash equilibrium.

Networks of Influence Diagrams (NIDs)

• Graphical language where each node is a Multi Agent Influence Diagram (MAID) (Koller and Milch, 2001)
• NIDs
  – Represent uncertainty over decision-making models.
  – Compute an equilibrium given agents’ beliefs, which might be incorrect !

Rock-Paper-Scissors (RoShamBo)

• In each round, players simultaneously choose one of rock, paper, or scissors.
• If the same item is chosen, result is tied.
• Otherwise, rock crushes paper, paper covers rock, and scissors cut paper.
• Game has a single mixed Nash equilibrium where both players randomize {1/3,1/3,1/3}, with expected utility zero.
1\textsuperscript{st} RoShamBo Competition (Billings 2000)

- Automatic RoShamBo players competed against each other playing multiple rounds.
- Nash equilibrium players broke even against any opponent. Opponent modelers did much better.
- Straightforward prediction will fail. Opponents disguise their strategy and attempt to counter model you (meta deliberators).
- Winning program - Iocaine powder (Egnor 2000)

Iocaine Powder’s reasoning

- Mary and John are playing rounds of RoShamBo. Suppose there exists a predictor $P$ that depends on prior history (ex. Predictor says \textit{paper})

Networks of Influence Diagrams

- A NID is a DAG where each node (or block) is a possible mental model.
- The root block (top-level) represents the real-world from the modeler’s point of view.
- Uncertainty over models is quantified by edge label.
Solving RoShamBo NID

- Compute a best-response for John at each round given his model description – distribution over $P$, meta-strategies and history.
- If we knew distribution over $P$ and Mary’s meta-strategies, we could compute John’s expected utility for \{rock, paper, scissors\} and pick the best move.
- … but these are unknown!
- Estimate them based on Mary’s play.

Solving NIDs

- Output a best-response strategy for each agent at each block, given his beliefs.
- We use a bottom-up solution algorithm
  - Eliminate each node by passing best-response of modeled decision to the parent.
  - See paper for details.

Possible solution

- Estimate missing parameters based on Mary’s play over time.

<table>
<thead>
<tr>
<th>Mary’s Play</th>
<th>Prob(Meta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>0.6</td>
</tr>
<tr>
<td>Paper</td>
<td>0.2</td>
</tr>
<tr>
<td>Scissors</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P</th>
<th>Prob(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>paper</td>
<td>0.2</td>
</tr>
<tr>
<td>rock</td>
<td>0.6</td>
</tr>
<tr>
<td>scissors</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prob(P)</th>
<th>Rock</th>
<th>Paper</th>
<th>Scissors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>0.5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>
On-line learning algorithm for NID parameters

- For any unknown random variable -
  - Assign arbitrary distribution to variable.
  - Compute a maximum-likelihood estimate for missing parameter based on observation using frequency counts or EM algorithm.
- We require that any missing variable exist in the real-world model.
- See paper for details.

Using the RoShambo NID in tournament

- We ran 10 games of 3,000 rounds each.
- RoShamBo NID results
  - Consistently beat “pattern” players.
  - Did not do well against meta deliberators.
- The symptom – meta strategies were not being learned. Uncertainty was folded into predictor $P$.

\[
\text{Observation}\quad \begin{cases} 
\text{Mary played Rock} & \text{Mary is following } P=\text{rock} \\
\text{Mary is playing } BR(BR(P=\text{paper})) & \text{other case}
\end{cases}
\]

The problem

- Meta-deliberators play a non-stationary strategy. At each round
  - Mary might be playing pattern.
  - Mary might be playing some meta-strategy.
- Define Mary’s “hyper” strategy as a lottery

\[
\begin{aligned}
\text{Prob. } q & \quad \text{Prob. } 1-q \\
\text{Follows } P & \quad \text{Mixed strategy over } P, BR(BR(P)) \quad BR(BR(BR(BR(P))))
\end{aligned}
\]

Learning non-stationary strategies

- Mary can play pattern (follows predictor $P$) or some meta-strategy.
- Let $h$ be some prior move history.
- Distribution over Mary’s hyper strategies depends on the frequency of $h$.
- First time we observe $h$, we assume Mary plays pattern. Otherwise, Mary is playing a meta-strategy with increasing probability that depends on $\#h$. 

## Example

<table>
<thead>
<tr>
<th>Prior move history</th>
<th>Mary’s play</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>pattern</td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>pattern</td>
<td></td>
</tr>
<tr>
<td>Scissors</td>
<td>BR(BR(P))</td>
<td></td>
</tr>
</tbody>
</table>

## Results

- Iocaine Powder NID consistently beat meta deliberators in the competition.
- It did not beat Iocaine Powder.

## Iocaine Powder vs. RoShamBo NID

- Why is it difficult to beat Iocaine Powder?
  - Erratically switches between meta strategies.
  - Considers the entire move history.
- NID did not explicitly model Iocaine Powder.

## Technical Details

- RoShamBo NID also includes predictor for Mary.
- NID syntax
  - Uncertainty over models is expressed by chance variables in the language.
Future Work

• Theoretical foundations.
  – NID semantics.
  – Relationship to Bayesian games.
• Modeling bounded rationality.
  – Multi-agent negotiation game between humans and computer players.
• Beating Iocaine Powder.

Conclusion

• NIDs are a solid foundation for opponent modeling.
  – Natural representation.
  – Offer a natural framework for learning.
• NIDs have many applications.
  – Opponent modeling.
  – Collusions and alliances.
• Come to plenary session talk on Wednesday