Safe Pareto improvements for delegated game playing
Caspar Oesterheld and Vincent Conitzer
{caspar,conitzer}@cs.duke.edu
Department of Computer Science, Duke University

Delegated game playing

<table>
<thead>
<tr>
<th></th>
<th>DM</th>
<th>RM</th>
<th>DL</th>
<th>RL</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM</td>
<td>-5,-5</td>
<td>2,0</td>
<td>5,-5</td>
<td>5,-5</td>
</tr>
<tr>
<td>RM</td>
<td>0,2</td>
<td>1,1</td>
<td>5,-5</td>
<td>5,-5</td>
</tr>
<tr>
<td>DL</td>
<td>-5,-5</td>
<td>-5,-5</td>
<td>1,1</td>
<td>2,0</td>
</tr>
<tr>
<td>RL</td>
<td>-5,-5</td>
<td>-5,-5</td>
<td>0,2</td>
<td>1,1</td>
</tr>
</tbody>
</table>

- Trusted representatives competent in strategic choice
  => Default: aligned delegation
- DL,RL are strictly dominated
  => never played
- **Equilibrium selection problem**
  => Pareto-suboptimal outcome (DM,DM) might occur

An unsafe approach

- **Publicly visible contracts** à la program equilibrium (Tennenholtz 2004) or mediated equilibrium (Monderer and Tennenholtz 2009)
- Each player's contract says: "Play RM if the other player adopts an analogous contract."
- Unclear whether this is good for both players.

A safe Pareto improvement

- Each player's contract says: "Assume this alternative payoff matrix if the other player adopts an analogous contract."
- The **new game** is essentially isomorphic to the original game.
- **Safe Pareto improvement** on the original game: outcome of new game is weakly better for both original players with certainty.

In the paper...
- Formal grounding
- Theorem: All safe Pareto improvements use equivalences between games.
- Theorem: Finding Safe Pareto improvements is NP-complete.