A Formal Separation Between Strategic and Nonstrategic Behavior

James R. Wright
Alberta Machine Intelligence Institute
University of Alberta

Joint work with Kevin Leyton-Brown (UBC)

Contributions:

1. A formal definition of strategic behavior that is not equivalent to perfect rationality

2. A constructive characterization that precisely distinguishes between strategic and nonstrategic behavior
Strategic vs. Nonstrategic Agents

- Behavioral game theorists often model **boundedly rational agents** (e.g., humans) using **iterative models** such as the **level-k model**:
  - **Level-0 agents**: behave uniformly at random
  - **Level-1 agents**: best respond to level-0 agents
  - **Level-2 agents**: best respond to level-1 agents

**Question**: What kinds of other behavior "count" as nonstrategic?

- Most work argues heuristically for certain rules
- (truthful reporting, largest number, etc.)

### Nonstrategic

<table>
<thead>
<tr>
<th>Uniform randomization</th>
<th>$a_i \in A_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxmin</td>
<td>$a_i \in \arg\max_{a_i} \min_{a_i} u_i(a_i, a_i)$</td>
</tr>
<tr>
<td>Maxmax</td>
<td>$a_i \in \arg\max_{a_i} \max_{a_i} u_i(a_i, a_i)$</td>
</tr>
<tr>
<td>Max total payoffs</td>
<td>$a_i \in \arg\max_{a_i} \max_{a_i} \sum_{a_i} u_i(a_i, a_i)$</td>
</tr>
<tr>
<td>Min unfairness*</td>
<td>$a_i \in \arg\min_{a_i} \min_{a_i} \left</td>
</tr>
<tr>
<td>Nash equilibrium</td>
<td>${a_i : \exists a_i : (a_i, a_i) \text{ is Nash equilibrium} }$</td>
</tr>
</tbody>
</table>
Strategic Agents

**Definition:** A behavioral model is **strategic** if it is both **other responsive** and **dominance responsive**.

**Definition:**
A behavioral model is **dominance responsive** if, for every pair of games where some action is strictly dominant in one game and strictly dominated in the other, the model does not behave identically:

\[ f_i(G) \neq f_i(G') \ \forall G, G' \text{ with } a_i^* \text{ dominant in } G \text{ and dominated in } G' \]

**Definition:**
A behavioral model is **other responsive** if there exists any pair of games that differ only in the payoffs of the other agents in which the model predicts different behavior:

\[ \exists G, G' : f_i(G) \neq f_i(G') \land \forall a \in A : u_i(a) = u_i'(a) \]

**Theorem:** All of QRE, Nash equilibrium, correlated equilibrium, cognitive hierarchy, and level-k(*) are (profiles of) **strategic** behavioral models.
Elementary Behavioral Models

Definition:
A behavioral model $f_i$ is **elementary** if it can be represented as $f_i(G) = h(\Phi(G))$, where:

- for all games $G = (N, A, u)$, for all $a \in A$, $\Phi(G)_a = \varphi(u(a))$,
- $\varphi$ satisfies **no smuggling**, and
- $h$ is an **arbitrary function** that maps $\mathbb{R}^A \rightarrow \Delta(A_i)$

**Main Theorem:** No **elementary** behavioral model $f_i$ is **strategic**.