Learning Robust Helpful Behaviors in Two-Player Cooperative Atari Environments

**Introduction.**
- We study the problem of learning helpful behavior: learning to cooperate with differently-skilled and diverse partners in the context of two-player, cooperative Atari games.
- We show robust performance of these Helper-AIs when paired with different kinds of partners (both human and artificial agents), including partners that they have not previously encountered during training.

**Helper-AI for Cooperative Atari 2600.**

**Intervention-AIs with Algs.**

<table>
<thead>
<tr>
<th>Two player Space Invaders</th>
<th>Partner Al</th>
<th>Two player Fall Down</th>
<th>Partner Al</th>
</tr>
</thead>
<tbody>
<tr>
<td>with self</td>
<td>808</td>
<td>1344</td>
<td>2144</td>
</tr>
<tr>
<td>with S5</td>
<td>1701</td>
<td>2346</td>
<td>3084</td>
</tr>
<tr>
<td>with Intervention Al (0.105 cost)</td>
<td>1462</td>
<td>1524</td>
<td>2005</td>
</tr>
<tr>
<td>with Intervention Al (0.075 cost)</td>
<td>1028</td>
<td>1224</td>
<td>2645</td>
</tr>
</tbody>
</table>

**Robust Helper-Al Behavior.**

1. Helpful behavior vs. expert behavior:
   - Pairing an agent with an expert-skill agent consistently reduces performance relative to self-pairing.
   - There is decisive and consistent performance improvement from pairing an AI with its on-target Helper-Al.

2. Robust helpful behavior:
   - There is a consistent improvement in performance when pairing an AI with an off-target Helper-Al compared to the performance from self-pairing.

3. Robust helpful behavior, bounded helpers:
   - The bounded-Helper-Al, bH(S5), provides a consistent improvement in performance for partner agents relative to self-pairing.

**Understanding Helper-Al Behavior.**

- Locations in two-player, Cooperative Space Invaders: Human subjects start at location 117 (at the right) and the Al start at location 35 (at the left).
- Helper-Als tend to spend less time at their initial location and play more in the center of the screen.
- Reasons for episode termination in two-player, Cooperative Space Invaders over 100 games, with partner AI S5, and varying the agent used in the role of Player 1.
- When Player 1 is replaced with Helper-Al H(S5), overall miscoordination goes down to 15%.

**Helper-Al Transfer to Human Partners.**

- Comparative performance in Cooperative Space Invaders when pairing Algs with S5 (another AI) or ten different human subjects.
- The decisive performance advantage of the Helper-Als, compared with pairing with either S5 or S4, holds up in transferring to this human environment.
- The bottom half of the table reports results for bH(S5) and the randomized-start position Helper-Als, H(S5), in a setting where the human subjects are sometimes randomly teleported to different positions and sometimes asked to do something unexpected for a period of time.

References:

Paul Tylikin
Goran Radanovic
David C. Parkes
petylikin@g.harvard.edu
gradanovic@mpi-sws.org
parkes@eecs.harvard.edu