**The Role of Intention Recognition in The Evolution of Cooperative Behavior**

Han The Anh*, Luis Moniz Pereira, Francisco C. Santos

Department of Informatics, Faculty of Science and Technology, New University of Lisbon, Portugal

Intention recognition is ubiquitous in most social interactions among Humans and other primates. However, the role of intention recognition in the emergence of cooperative actions remains elusive. Using the tools of evolutionary game theory, here we address, for the first time, the role played by intention recognition in the final outcome of cooperation in large populations of self-regarding individuals. By equipping individuals with the capacity of assessing intentions of others in the course of repeated Prisoner's Dilemma interactions, we show how intention recognition opens a window of opportunity for cooperation to thrive, as it precludes the invasion of pure cooperators by random drift while remaining robust against defective strategies. Intention recognizers are able to assign an intention to the action of their opponents based on an acquired corpus of possible intentions. We show how intention recognizers can prevail against most famous strategies of repeated dilemmas of cooperation, even in the presence of errors.

Iterated Prisoner's Dilemma

Payoff matrix

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>R</td>
<td>S</td>
</tr>
<tr>
<td>D</td>
<td>T</td>
<td>P</td>
</tr>
</tbody>
</table>

C: cooperate; D: defects
R: reward; S: sucker
T: temptation; P: punishment

T > R > P > S; 2R > T + S

In a single round, it is always better to defect. Cooperation may be rewarded if the game is iterated.

Strategies in iterated PD (intentions to be recognized)

TFT: plays what opponent chose in previous round.
WSLS: win stays (with the same move), loses shift (to different move).
AIC: always cooperates.
TFT: always defects.

Intention-based decision making

Intention recognizer plays what would provide the greatest expected payoff against the recognized strategy

$$\theta(s, I) = \arg \max_{\xi \in \{C, D\}} E[\xi]$$

s: current game state
I: opponent's recognized intention

Donor game: donor pays a cost c for the recipient to obtain a benefit b.

Comparing IR with WSLS and TFT when interacting with AIC and AID

Intention recognizers prevail for different level of noise and benefit-to-cost ratios.

Evolutionary simulations with intention recognition

Competition of several strategies

IR wins the competition. AID wins in absence of IR.

References.


*han@fct.unl.pt

Han&Pereira&Santos. IJCAI’2011, Barcelona