
Learning to recognize intentions resolves cooperation dilemmas

The Anh Han

H.ANH@AI.VUB.AC.BE

Artificial Intelligence lab, Vrije Universiteit Brussel & MLG, Université Libre de Bruxelles

Luís Moniz Pereira

LMP@FCT.UNL.PT

Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

Tom Lenaerts

TOM.LENAERTS@ULB.AC.BE

MLG, Université Libre de Bruxelles & Artificial Intelligence lab, Vrije Universiteit Brussel

Francisco C. Santos

FRANCISCOCSANTOS@IST.UTL.PT

INESC-ID & Instituto Superior Técnico, Universidade de Lisboa

Keywords: Intention recognition, Social dilemma, Evolution of Cooperation, Bayesian Network.

Few problems have created the combined interest of so many unrelated areas as the evolution of cooperation (Nowak, 2006; Sigmund, 2010). As a result, several mechanisms have been identified to work as catalyzers of cooperative behavior (see survey in (Nowak, 2006; Sigmund, 2010)). Yet, these studies, mostly grounded on evolutionary dynamics and game theory, have neglected the important role played by intention recognition (Han & Pereira, 2013) in behavioral evolution. We discuss here our work in (Han et al., 2011; Han et al., 2012), wherein we explicitly studied the role of intention recognition in the evolution of cooperative behavior. By equipping some individuals with the capacity of learning the intentions of others in the course of a prototypical dilemma of cooperation—the repeated prisoner’s dilemma—we show how intention recognition is favored by natural selection, opening a window of opportunity for cooperation to thrive. We introduced new strategies that are able to learn to predict an opponent’s intention, either on the basis of an acquired corpus consisting of possible plans achieving that intention (Han et al., 2012), or through incrementally constructing a Bayesian Network model that assesses the internal dynamics of trust between intention recognizers and their opponents (Han et al., 2011). In both cases, the probability that opponents have a certain intention is computed conditional on their past actions and the recognizers’ actions, and intention assignment is based on comparing conceivable intentions’ probabilities. Our results show that intention recognizers prevail against the most successful strategies in the context of the iterated Prisoner’s Dilemma (e.g. tit-for-tat like strategies and win-stay-lose-shift (Sigmund, 2010)), and promote a significantly high level of cooperation, even in the presence of noise and re-

duction of fitness associated with a cognitive cost for performing intention recognition. The success of intention recognizers is grounded either on the ability to learn to exploit unconditional cooperators while remaining robust against unconditional defectors (Han et al., 2012); or on the ability to learn to deal with noise through assessing the opponent’s real intention (Han et al., 2011).

Overall, our study offers new insights on the complexity and beauty of behavioral evolution driven by elementary forms of cognition and learning ability.

Acknowledgments

TAH acknowledges the support from the FWO. TL acknowledges the support from the FWO and FRNS.

References

- Han, T. A., & Pereira, L. M. (2013). State-of-the-art of intention recognition and its use in decision making. *AI Communication*, 26, 237–246.
- Han, T. A., Pereira, L. M., & Santos, F. C. (2011). Intention recognition promotes the emergence of cooperation. *Adaptive Behavior*, 19, 264–279.
- Han, T. A., Pereira, L. M., & Santos, F. C. (2012). Corpus-based intention recognition in cooperation dilemmas. *Artificial Life journal*, 18, 365–383.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314, 1560.
- Sigmund, K. (2010). *The calculus of selfishness*. Princeton University Press.