The Role of Intention Recognition in The Evolution of Cooperative Behavior

Han The Anh and **Luís Moniz Pereira** and **Francisco C. Santos**
Centro de Inteligência Artificial (CENTRIA)
Departamento de Informática, Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
$h.anh@fct.unl.pt$, $lmp@di.fct.unl.pt$, $fcsantos@fct.unl.pt$

Abstract

Given its ubiquity, scale and complexity, few problems have created the combined interest of so many unrelated areas as the evolution of cooperation. 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. Our approach invites the adoption of other classification and pattern recognition mechanisms common among Humans, to unveil the evolution of complex cognitive processes in the context of social dilemmas.

1 Introduction

In multi-agent systems, the problem of intention recognition appears to be crucial whenever the achievement of a goal by an agent does not depend uniquely on its own actions, but also on the decisions of others. This is particularly common when agents cooperate or have to coordinate their actions to achieve a task, especially when the possibility of communication may be limited [Heinze, 2003; Segbroeck *et al.*, 2010]. For example, in heterogeneous agent systems it is likely that agents speak different languages, have different designs or different levels of intelligence; hence, intention recognition may be the only way agents understand each other to secure successful cooperation. Moreover, agents often attempt to hide their real intentions and make others believe in pretense ones [Pereira and Han, 2011].

Intention recognition is defined, in general terms, as the process of becoming aware of another agent's intention and, more technically, as the problem of inferring an agent's intention through its actions and their effects on the environment [Heinze, 2003].

The problem of intention recognition has been paid much attention in AI, Philosophy and Psychology for several decades [Kautz and Allen, 1986; Charniak and Goldman, 1993; Bratman, 1987; Geib and Goldman, 2009]. Whereas intention recognition has been extensively studied in small scale interactive settings, there is an absolute lack of modelling research with respect to large scale social contexts; namely the evolutionary roles and aspects of intention recognition.

In this work, we study the role of intention recognition for one of the most challenging but intriguing issues, traversing areas as diverse as Biology, Economics, Artificial Intelligence, Political Science, or Psychology: the problem of *evolution of cooperation* [Axelrod, 1984]. Why would natural selection equip selfish individuals with altruistic tendencies while it incites competition between individuals and thus apparently rewards only selfish behavior? Several mechanisms responsible for promoting cooperative behavior have been recently identified [Sigmund, 2010], including kin selection, direct and indirect reciprocity, network reciprocity, group selection (see [Nowak, 2006] for a survey). Here we wish to understand how cooperation emerges from the interplay between population dynamics and individuals' cognitive abilities, namely the ability to perform intention recognition. Like addressing the problem of evolution of cooperation, our study is carried out within the framework of Evolutionary Game Theory (EGT) [Hofbauer and Sigmund, 1998]. Here, individual success (or fitness) is expressed in terms of the outcome of a 2-person game, which, in turn, is used by individuals to copy others whenever these appear to be more successful.

Intention recognition can be found abundantly in many kinds of interactions and communications, not only in Human but also many other species [Tomasello, 2008]. The knowledge about intention of others in a situation could enable to plan in advance, either to secure a successful cooperation or to deal with potential hostile behaviours [Pereira and Han, 2011]. Given the advantage of knowing others' intention and the abundance of intention recognition in many different species, we believe that intention recognition should be taken into account to design appropriate strategies—in studying and in explaining all the mechanisms for the evolution of cooperation.

In this work we take a first step towards employing intention recognition within the framework of repeated interactions. Similarly to direct reciprocity [Trivers, 1971] intention recognition is being performed using the information about past *direct* interactions. As usual, the inputs of an intention recognition system are a set of conceivable intentions and a set of plans achieving each intention (plan library [Geib and Goldman, 2009] or plan corpus [Blaylock and Allen, 2003]). In this EGT context, conceivable intentions are the strategies already known to the intention recognizer, whose recognition model is learnt from a plan corpus consisting of sequences of moves (called plan sessions) of different strategies. There have been several successful corpus-based intention recognition models in the literature [Blaylock and Allen, 2003; Blaylock and Allen, 2004; Armentano and Amandi, 2009], and we adjust one to the current work in lieu of supplying a novel one (see Subsection 3.1). The rationale of the corpus-based approach firstly relies on the idea of nature-nurture co-evolution or experience inheritance [Richerson and Boyd, 2006; Shennan, 2002]: the corpus represents ancestors' experience in interacting with known strategies. Additionally, intention recognizers can use themselves as a framework for learning and understanding those strategies by self-experimenting them—as suggested by the famous 'like-me' framework [Meltzoff, 2007]. This is often addressed in the context of the "Theory of Mind" theory, neurologically relying in part on "mirror neurons", at several cortical levels, as supporting evidence [Iacoboni *et al.*, 2005]. In addition, we offer a method to acquire a *rational* decision making model from the plan corpus, that states what to play with a co-player based on the recognized intention and the game's current state. Rationality means the decision maker attempts to achieve the greatest expected benefit for himself/herself. The model is discussed in Subsection 3.2.

2 EGT Plan Corpus

2.1 Interaction between Agents

Interactions are modeled as symmetric two-player games defined by the payoff matrix

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R, R & S, T \\ T, S & P, P \end{pmatrix} \end{array}$$

A player who chooses to cooperate (C) with someone who defects (D) receives the sucker's payoff S, whereas the defecting player gains the temptation to defect, T. Mutual cooperation (resp., defection) yields the reward R (resp., punishment P) for both players. Depending on the ordering of these four payoffs, different social dilemmas arise [Macy and Flache, 2002; Santos *et al.*, 2006; Sigmund, 2010]. Namely, in this work we are concerned with the Prisoner's Dilemma (PD), where $T > R > P > S$. In a single round, it is always best to defect, but cooperation may be rewarded if the game is repeated. In repeated PD, it is also required that mutual cooperation is preferred over an equal probability of unilateral cooperation and defection ($2R > T + S$); otherwise alternating between cooperation and defection would lead to a higher payoff than mutual cooperation.

The repeated PD is usually known as a story of tit-for-tat (TFT), which won both Axelrod's tournaments [Axelrod, 1984; Axelrod and Hamilton, 1981]. *TFT* starts by cooperating, and does whatever the opponent did in the previous round. It will cooperate if the opponent cooperated, and will defect if the opponent defected. But if there are erroneous moves because of noise (i.e. an intended move is wrongly performed with a given execution error, referred here as "noise"), the performance of *TFT* declines, in two ways: (i) it cannot correct errors and (ii) a population of *TFT* players is undermined by random drift when *AllC* (always cooperate) mutants appear (which allows exploiters to grow). Tit-for-tat is then replaced by generous tit-for-tat (GTFT), a strategy that cooperates if the opponent cooperated in the previous round, but sometimes cooperates even if the opponent defected (with a fixed probability $p > 0$) [Nowak and Sigmund, 1992]. *GTFT* can correct mistakes, but remains suffering the random drift.

Subsequently, *TFT* and *GTFT* were replaced by win-stay-lose-shift (WSLS) as the winning strategy chosen by evolution [Nowak and Sigmund, 1993]. *WSLS* repeats the previous move whenever it did well, but changes otherwise. *WSLS* corrects mistakes better than *GTFT* and does not suffer random drift. However, it is exploited by pure defectors *AllD*.

In the following, abusing notations, R, S, T and P are also referred to as game states (in a single round or interaction). We also use E (standing for *empty*) to refer to the game state having no interaction.

2.2 Plan Corpus Description

We made an assumption that all strategies to be recognized have the memory size bounded-up by M ($M \geq 0$)—i.e. their decision at the current round is independent of the past rounds that are at a time distance greater than M.

An action in the corpus is of the form $s_1 ... s_M \xi$, where $s_i \in \{E, R, T, S, P\}$, $1 \leq i \leq M$, are the states of the M last interactions, and $\xi \in \{C, D\}$ is the current move. We denote by Σ_M the set of all possible types of action. For example, $\Sigma_1 = \{EC, RC, TC, SC, PC, ED, RD, TD, SD, PD\}$. EC and ED only occur in the first round, when the move is C and D, respectively (the state of the round before the first one is E).

This way of encoding actions and the assumption about the players' memory size lead to the equivalent assumption that the action in the current round is independent of the ones in previous rounds, regardless of the memory size. Furthermore, this encoding method enables to save the game states without having to save the co-player's moves, thus simplifying the corpus representation, described below.

Let us suppose we have a set of strategies to be recognized. The plan corpus for this set consists of a set of plan sessions generated for each strategy in the set. A plan session of a strategy is a sequence of actions played by that strategy (more precisely, a player using that strategy) against another player.

3 Intention Recognizers' Models

3.1 Corpus-based Intention Recognition Model

We can use any corpus-based intention recognition model in the literature for this work. The most successful works are

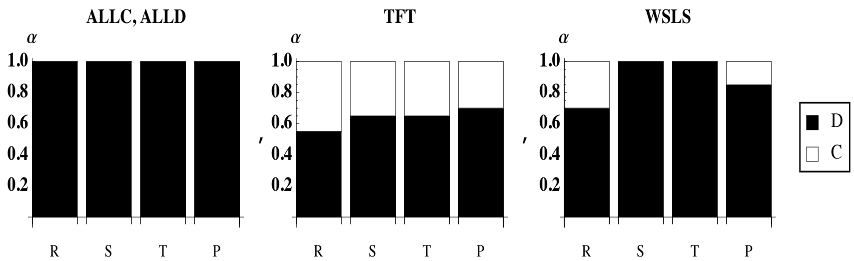

Figure 1: Decision making model for different values of α. If the recognized intention is *AllC* or *AllD*, intention recognizers (IR) always defect, regardless of the current states. If it is *TFT*, *IR* cooperates when α is large enough, regardless of the current states. If it is *WSLS*, if the current states are S or T, IR always defects; otherwise, *IR* cooperates for large enough α. This model is acquired for a PD with $R = 1, S = -1, T = 2, P = 0$. The model has the same behavior for all PD payoff matrixes used in this paper.

described in [Blaylock and Allen, 2003; Blaylock and Allen, 2004; Armentano and Amandi, 2009]. Blaylock and Allen use the bigram statistical model, making the assumption that the current action only depends on the previous one. Armentano and Amandi attempts to avoid this assumption by using the Variable-Order Markov Model. In our work, because of the independence of actions, we can derive an even simpler model than that of Blaylock and Allen, as described below.

Let I_i, $1 \le i \le n$, be the intentions to be recognized, and $O = \{A_1, ..., A_m\}$ the set of current observed actions. The intention recognition task is to find the most likely intention $I^\star \in \{I_1, ..., I_n\}$ given the current observed actions, i.e.

$$I^\star = \arg\max_{I_i:1\le i\le n} P(I_i|A_1, ..., A_m)$$
$$= \arg\max_{I_i:1\le i\le n} \frac{P(I_i)\prod_{j=1}^{m} P(A_j|I_i, A_1, ..., A_{j-1})}{P(O)}$$

The second equation is obtained by applying the Bayes' and then Chain rules. Since the denominator $P(O)$ is a positive constant, we can ignore it. Then, because of the independency amongst actions, we obtain

$$I^\star = \arg\max_{I_i:1\le i\le n} P(I_i) \prod_{j=1}^{m} P(A_j|I_i) \quad (1)$$

Note that this model is independent of the memory size M. Also note that if two intentions are assessed with the same probability, the model predicts the one with higher priority; these are set depending on the behavioral attitude of the intention recognizer. For example, in Figure 2, if *IR*'s co-player cooperates in the first round, the co-player can be predicted as either *AllC*, *WSLS* or *TFT*. Being concerned of *TFT*'s and *WSLS*'s retaliation after a defect, *WSLS* and *TFT* should have higher priority than *AllC*.

3.2 Decision Making Model

We describe how to acquire a decision making model for an intention recognizer from the plan corpus. As a rational agent, the intention recognizer chooses to play what would provide it the greatest expected payoff against the recognized strategy (intention). Namely, from training data we need to extract the function $\theta(s, I)$:

$$\theta : \{E, R, T, S, P\}^M \times \{I_1, ..., I_n\} \to \{C, D\}$$

deciding what to play (C or D) given the current state $s = s_1...s_M$, where $s_i \in \{E, R, T, S, P\}$ $(1 \le i \le M)$, and the recognized intention $I \in \{I_1, ..., I_n\}$. It is done as follows. From the training plan sessions for each intention we compute the (per-round) average payoff the intention recognizer would receive with respect to each choice (C or D), for each possible state s. The choice giving greater payoff is chosen. Formally, let *DS(I)* be the set of all sequences of actions (plan sessions), $Sq = A_1....A_k$ ($A_i \in \Sigma_M$, $1 \le i \le k$), generated for intention I in the corpus and $\pi(Sq, j)$ the payoff the co-player of I gets at round j. In the following, if the sequence in which the payoff being computed is clear from the context, we ignore it and simply write $\pi(j)$. Thus,

$$\theta(s, I) = \arg\max_{\xi\in\{C,D\}} \Pi_{\alpha\xi}/N_{\alpha\xi} \quad (2)$$

where $\Pi_{\alpha\xi} = \sum_{A_1...A_k\in DS(I)} \sum_{A_i=s\xi} \sum_{j=i}^{k} \alpha^{j-i}\pi(j)$; $N_{\alpha\xi} = \sum_{A_1...A_k\in DS(I)} \sum_{A_i=s\xi} \sum_{j=i}^{k} \alpha^{j-i}$; and a further round's payoff is weighted less than a nearer round's payoff by a discounting factor $1/\alpha$ with $0 < \alpha \le 1$.

Note that, at the first round, there is no information about the co-player. The intention recognizer cooperates, i.e. $\theta(E^M, I) = C \;\; \forall I \in \{I_1, ..., I_n\}$.

Experimental results obtained from this model are provided in Subsection 4.3 (Figure 1).

4 Experiments and Results

4.1 Plan Corpus Generation

Let us start by generating a plan corpus of four of the most famous strategies within the framework of repeated games of cooperation: *AllC* (always cooperate), $AllD$ (always defect), *TFT* and *WSLS* (see above). Not only these strategies constitute the most used corpus of strategies used in this context, as most other strategies can be seen as a high-level composition of the principles enclosed in these strategies. Hence, intention recognizers map their opponent's behaviors to the closest strategy that they know and interact accordingly. When their knowledge is extended to incorporate new strategies, the models can be revised on the fly. However, here we do not deal with this issue.

We collect plan sessions of each strategy by playing a random choice (C or D) in each round with it. To be more thorough, we can also play all the possible combinations for each given number of rounds to be played. For example, if it is 10, there will be 1024 (2^{10}) combinations—C or D in each

IR:	C D D D D ...	**IR:**	C C D D D...	**IR:**	C C D C C...
AllD:	D D D D D ...	**AllC:**	C C C C C...	**IR:**	C C D C C...
IR:	C C D D C C C C...			**IR:**	C C D D C C C C...
TFT:	C C C D D C C C...			**WSLS:**	C C C D C C C C...

Figure 2: Interactions of *IR* with *AllC*, *AllD, TFT, WSLS* and another *IR*, in the absence of noise and $\alpha = 1$.

	AllC	AllD	TFT	WSLS	Total
Precision	0.859	**0.999**	0.818	0.579	**0.824**
Recall	0.859	0.999	0.818	0.575	0.824
Converg.	0.859	0.999	0.719	0.579	0.805

Table 1: Intention recognition results for each strategy and the total.

round. When noise is present, each combination is played repeatedly several times.

The training corpus is generated by playing with each strategy all the possible combinations 20 times, and for each number of rounds from 5 to 10. The testing dataset is generated by playing a random choice with each strategy in each round, and also for each number of rounds from 5 to 10. We continue until obtaining the same number of sessions as in the training dataset (corpus). Both datasets are generated in presence of noise (namely, an intended move is wrongly performed with probability 0.05).

4.2 Intention Recognition Model Evaluation

Evaluation Metrics

For evaluating the intention recognition model, we use three different metrics. *Precision* and *recall* report the number of correct predictions divided by total predictions and total prediction opportunities, respectively. If the intention recognizer always makes a prediction (whenever it has the opportunity), recall is equal to precision. *Convergence* is a metric that indicates how much time the recognizer took to converge on what the current user goal/intention was. Formal definitions of the metrics can be found in [Armentano and Amandi, 2009].

Results

The intention recognition model is acquired using the training corpus. Table 1 shows the recognition results of the model for the testing dataset, using the three metrics described above. We show the recognition result for each strategy, and for the whole dataset. Given that the training as well as the testing datasets are generated in presence of noise, the achieved intention recognition performance is quite good. In the next section, we study the performance of players using this intention recognition model (called *IR*) in the large scale population setting: what is the role of intention recognition for the emergence of cooperation?

4.3 Decision Making Model Acquisition

The decision making model (in Subsection 3.2) is acquired using the training corpus (Figure 1). Figure 2 describes how *IR* interacts with other strategies. Except with *AllD*, *IR* plays C in the first two rounds with other strategies: IR always plays C in the first round, and since others also play C (thus, the action is EC), they are predicted as a *TFT* (since $P(EC|ALLC) = P(EC|TFT) = P(EC|WSLS) \gg P(EC|AllD)$)—therefore, *IR* plays C in the second round. Note that here TFT has a higher priority than *WSLS*, which has a higher priority than *AllC*. In the third round, these strategies are all predicted as *AllC* since they play C in the second round (and since $P(RC|ALLC) > P(RC|WSLS) > P(RC|TFT)$). Hence, *IR* plays D in this round. The moves of these strategies (the other *IR* plays D, others play C) classifies *IR* to be *WSLS*, and the other three remain to be *AllC* (since $P(RD|WSLS) > P(RD|TFT) \gg P(RD|AllC)$). The two inequalities $P(RC|WSLS) > P(RC|TFT)$ and $P(RD|WSLS) > P(RD|TFT)$, for big enough training corpus, are easily seen to hold: although *TFT* and *WSLS* equally likely play C (resp., D) after R, since *WSLS* corrects mistakes better than *TFT*, mutual cooperations are more frequent in plan sessions for *WSLS*. The reaction in the fourth round classifies *TFT* to be *TFT*, *IR* and *WSLS* to be *WSLS*, and *AllC* to be *AllC*; and like that in the subsequent rounds. From the fifth round on, *IR* cooperates with *WSLS*, *TFT* and another *IR*. If the number of rounds to be played is very large, up to some big round, these three strategies will be recognized as *AllC* again (since $P(RC|ALLC) > P(RC|WSLS) > P(RC|TFT)$), then the process repeats as from third round. In our corpus, it only happens after more than 100 rounds.

4.4 Evolutionary Dynamics: Analysis

Consider a population of *AllC*, *AllD* and *IR* players. They play the repeated PD. Suppose m ($m < 100$) is the average number of rounds. In absence of noise, the payoff matrix of *AllC*, *AllD* and *IR* in m rounds is given by (Figure 2, $\alpha = 1$)

$$
\begin{array}{cc}
 & \begin{array}{ccc} \mathit{AllC} & \mathit{AllD} & IR \end{array} \\
\begin{array}{c} \mathit{AllC} \\ \mathit{AllD} \\ IR \end{array} &
\begin{pmatrix}
Rm & Sm & 2R + S(m-2) \\
Tm & Pm & T + P(m-1) \\
T(m-2) + 2R & P(m-1) + S & R(m-1) + P
\end{pmatrix}
\end{array}
$$

In each round, *AllC* cooperates. Thus, its co-player would obtain a reward R if it cooperates and a temptation payoff T otherwise; Hence, in playing with *AllC* (first column of the matrix), another *AllC* obtains m times of R since it cooperates in each round; *AllD* obtains m times of T since it defects in each round; and IR obtains 2 times of R and $(m-2)$ times of T since it cooperates with $AllC$ in the first two rounds and defects in the remaining rounds (Figure 2). Other elements of the matrix are computed similarly.

Pairwise comparisons [Nowak, 2006] of the three strategies lead to the conclusions that *AllC* is dominated by *IR* and that *IR* is an evolutionary stable strategy if $R(m-1) > T + P(m-2)$, which always holds for $m \geq 3$. Evolutionarily stable strategy is a strategy which, if adopted by a

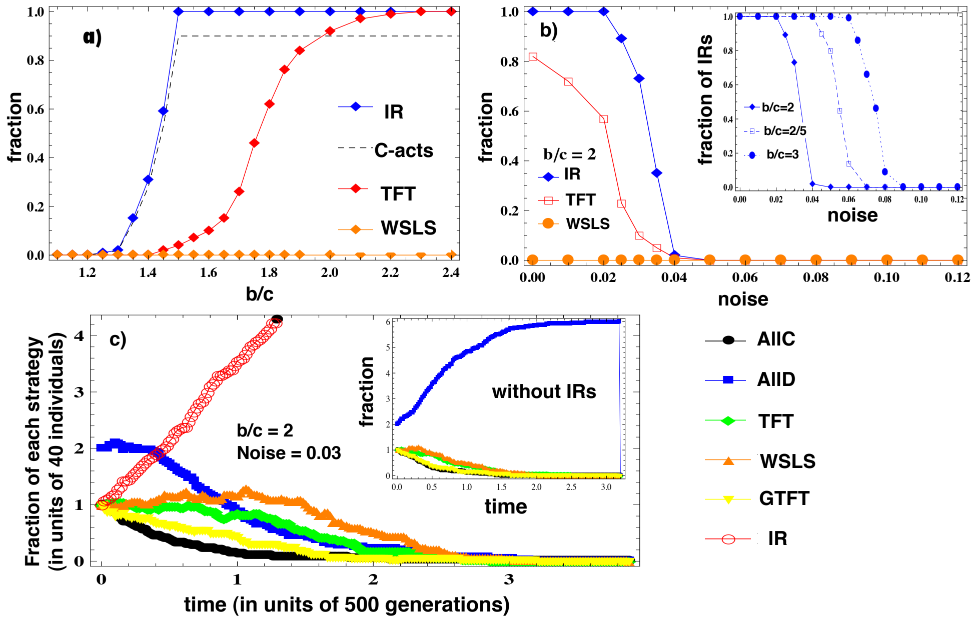

Figure 3: Simulation results for Donor game.

population of players, cannot be invaded by any alternative strategy that is initially rare [Hofbauer and Sigmund, 1998]. This condition guarantees that once *IR* dominates the population, it becomes stable (for $m \geq 3$).

4.5 Evolutionary Simulations

In presence of noise, it was not easy to provide an exact mathematical analysis. Instead, we will study this case using computer simulations. For convenience and a clear representation of simulation results, we use Donor game [Sigmund, 2010]—a famous special case of the PD—where $T = b,\ R = b - c,\ P = 0,\ S = -c$, satisfying that $b > c > 0$.

We start with a well-mixed population of size N, whose individuals use different strategies. In each round of a generation, each individual interacts with all others, engaging in a PD game. As usual, the accumulated payoff from all interactions emulates the individual *fitness* (f) or social *success* and the most successful individuals will tend to be imitated by others. Such behavioral evolution is implemented using the pairwise comparison rule [Traulsen *et al.*, 2006]: after each generation an individual A will adopt the strategy of a randomly chosen individual B with a probability given by the Fermi function (from statistical physics) $p(A, B) = (1 + e^{-\beta(f(B)-f(A))})^{-1}$. The quantity β, which in physics corresponds to an inverse temperature, controls the intensity of selection.

The results are shown in Figure 3. All results were obtained averaging over 100 runs, for $m = 10$, $N = 100$ and $\beta = 0.1$. In (a) and (b), we consider populations of three strategies—*AllC*, *AllD* and either *IR*, *TFT* or WLSL—equally distributed at the beginning. We plot the final fraction of *IR*, *TFT* and *WSLS*. All simulations end up in a homogeneous state (i.e. having only one type of strategy) in less than 5000 generations. Our results show that *IR* prevails *TFT* and *WSLS* for different benefit-to-cost ratios b/c (panel a) and for different levels of noise (panel b). For a small ratio b/c (around 1.2), *IR* starts having winning opportunity, and from around 1.4 the population always converges to the homogeneous state of *IR*. For *TFT*, they are 1.4 and 2.1, respectively. *WSLS* has no chance to win for $b/c \leq 2.4$. The dashed black curve in (a) shows that the fraction of cooperation in the population of *AllC*, *AllD* and *IR* is monotonic to b/c. In (b), our results show that, in the presence of noise, *IR* is advantageous when compared with *TFT* and *WSLS*. This result is robust to chances on the value of b/c (the inset of panel b)). In (c), we consider a more complex setting where the population consists of several types of strategies: *AllC, AllD, TFT, WSLS, GTFT* (probability of forgiveness of a defect $p = 0.5$) and *IR*. Except for the defective *AllD* and rational *IR*, the other strategies are cooperative. Thus, instead of initially being equally distributed, we include a higher fraction of *AllD*s in the initial population. Namely, each type has 40 individuals, and *AllD* has 80. *IR* always wins (main panel in c)). However, if *IR* individuals are removed, *AllD* is the winner (the inset figure), showing how *IR*s work as a catalyzer for cooperation. We have tested and obtained similar results for larger population sizes. Finally, in (a) and (b) we show how *WSLS* performs badly, as *WSLS* needs *TFT*s as a catalyst to perform well [Sigmund, 2010]—which can be observed in (c).

5 Concluding Remarks

The main contribution of this work is that we have shown, for the first time, that intention recognition promotes the emergence of cooperation. Given the broad spectrum of problems which are addressed using this cooperative metaphor, our result indicates how intention recognition can be pivotal in social dynamics. We have shown that a population with some fraction of intention recognizers acting rationally can lead to a stable cooperation. The intention recognition strategy has a greater range of benefit-to-cost ratios leading to cooperation than the most successful existent strategies (*TFT*, *WSLS*).

Secondly, our approach of using plan corpus makes a case for different other AI techniques to work with the problem of cooperation. In this work, we studied the role of intention recognition for the emergence of cooperation, but other cognitive abilities are also of great interest and importance, for example pattern recognition algorithms. Classification algorithms (or supervised learning in general) are clearly a good candidate. Indeed, intention recognition can be considered as a classification problem: the sequence of observed actions is classified into a known strategy. Clustering algorithms (or unsupervised learning in general) can be used to categorize the sequences of actions that are not fit with the known strategies. This is a way to learn about unknown strategies, categorize them, revise the model to take them into account (and pass the revised model to the successors).

Moreover, for the intention recognition community, given the rich set of strategies in the literature [Hofbauer and Sigmund, 1998; Sigmund, 2010], we have provided an important, easily extendable benchmark for evaluating intention recognition methods; especially as it is known that there are only one or two regularly used plan corpora available: the Linux Plan Corpus and its ancestor Unix Plan Corpus [Blaylock and Allen, 2003], and both are of a quite small size. Lastly, once we have an *IR* model, we can next employ it to enrich the corpus, and then iteratively acquire a new generation *IR* model, and so forth.

Acknowledgments

We thank the anonymous reviews for their comments. HTA (reference SFRH/BD/62373/2009) and FCS acknowledge the financial support of FCT-Portugal.

References

[Armentano and Amandi, 2009] Marcelo G. Armentano and Analía Amandi. Goal recognition with variable-order markov models. In *IJCAI'09*, pages 1635–1640, 2009.

[Axelrod and Hamilton, 1981] Robert Axelrod and W.D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[Axelrod, 1984] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, ISBN 0-465-02122-2, 1984.

[Blaylock and Allen, 2003] N. Blaylock and J. Allen. Corpus-based, statistical goal recognition. In *IJCAI'03*, pages 1303–1308, 2003.

[Blaylock and Allen, 2004] N. Blaylock and J. Allen. Statistical goal parameter recognition. In *ICAPS'04*, pages 297–304. AAAI, 2004.

[Bratman, 1987] M. E. Bratman. *Intention, Plans, and Practical Reason*. The David Hume Series, CSLI, 1987.

[Charniak and Goldman, 1993] E. Charniak and R.P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, 64(1):53–79, 1993.

[Geib and Goldman, 2009] C.W. Geib and R.P. Goldman. A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence*, 173:1101–1132, 2009.

[Heinze, 2003] C. Heinze. *Modeling Intention Recognition for Intelligent Agent Systems*. PhD thesis, The University of Melbourne, Australia, 2003.

[Hofbauer and Sigmund, 1998] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[Iacoboni *et al.*, 2005] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. C. Mazziotta, and G. Rizzolatti. Grasping the intentions of others with ones own mirror neuron system. *PLoS Biol*, 3(3):e79, 2005.

[Kautz and Allen, 1986] H. Kautz and J.F. Allen. Generalized plan recognition. In *AAAI'86*, pages 32–38, 1986.

[Macy and Flache, 2002] M. W. Macy and A. Flache. Learning dynamics in social dilemmas. *Proc. Natl. Acad. Sci. USA*, 99:7229–7236, 2002.

[Meltzoff, 2007] A. N. Meltzoff. The framework for recognizing and becoming an intentional agent. *Acta Psychol Amst*, 124(1):26–43, 2007.

[Nowak and Sigmund, 1992] Martin A. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355:250–253, 1992.

[Nowak and Sigmund, 1993] Martin A. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit for tat in prisoners dilemma. *Nature*, 364:56–58, 1993.

[Nowak, 2006] Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560, 2006.

[Pereira and Han, 2011] L. M. Pereira and T. A. Han. Intention recognition with evolution prospection and causal bayesian networks. In *Computational Intelligence for Engineering Systems*, pages 1–33. Springer, 2011.

[Richerson and Boyd, 2006] Peter J. Richerson and Robert Boyd. *Not By Genes Alone: How Culture Transforms Human Evolution*. The University of Chicago Press, 2006.

[Santos *et al.*, 2006] Francisco C. Santos, Jorge M. Pacheco, and Tom Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. USA*, 103:3490–3494, 2006.

[Segbroeck *et al.*, 2010] S. V. Segbroeck, S. de Jong, A. Nowe, F. C. Santos, and T. Lenaerts. Learning to coordinate in complex networks. *Adaptive Behavior*, 18:416–427, 2010.

[Shennan, 2002] Stephen Shennan. *Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution*. London: Thames and Hudson Ltd., 2002.

[Sigmund, 2010] Karl Sigmund. *The Calculus of Selfishness*. Princeton U. Press, 2010.

[Tomasello, 2008] M. Tomasello. *Origins of Human Communication*. MIT Press, 2008.

[Traulsen *et al.*, 2006] A. Traulsen, M. A. Nowak, and J. M. Pacheco. Stochastic dynamics of invasion and fixation. *Phys Rev E*, 74:011909, 2006.

[Trivers, 1971] R. L. Trivers. The evolution of reciprocal altruism. *Quaterly Review of Biology*, 46:35–57, 1971.