

# Corpus-based Intention Recognition in Cooperation Dilemmas

The Anh Han<sup>1</sup>, Luís Moniz Pereira<sup>1\*</sup>, Francisco C. Santos<sup>1,2,3</sup>

<sup>1</sup>Centro de Inteligência Artificial (CENTRIA), Departamento de Informática,  
Faculdade de Ciências e Tecnologia,  
Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

<sup>2</sup>DEI & INESC-ID, Instituto Superior Técnico, TU Lisbon,  
Taguspark, 2744-016 Porto Salvo, Portugal

<sup>3</sup>ATP-group, CMAF, Instituto para a Investigação Interdisciplinar,  
P-1649-003 Lisboa Codex, Portugal

emails: h.anh@campus.fct.unl.pt, lmp@fct.unl.pt, franciscocsantos@ist.utl.pt

---

\*Corresponding author. Email address: lmp@fct.unl.pt; Phone number (office): +351 212948533; Mobile number: +351 936778968; Fax: +351 212948541

## Abstract

Intention recognition is ubiquitous in most social interactions among humans and other primates. Despite this, the role of intention recognition in the emergence of cooperative actions remains elusive. Resorting to the tools of evolutionary game theory, herein we describe a computational model showing how intention recognition co-evolves with cooperation in populations of self-regarding individuals. By equipping some individuals with the capacity of assessing 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 introduce a new strategy (IR) that is able to assign an intention to the actions of opponents, on the basis of an acquired corpus consisting of possible plans achieving that intention, as well as to then make decisions on the basis of such recognized intentions. The success of *IR* is grounded on the unshameful exploitation of unconditional cooperators whilst remaining robust against unconditional defectors. In addition, we show how intention recognizers do indeed prevail against the most famous successful strategies of iterated dilemmas of cooperation, even in the presence of errors and reduction of fitness associated with a small cognitive cost for performing intention recognition.

**Keywords:** Evolution of Cooperation, Corpus-based Intention Recognition, Evolutionary Game Theory, Prisoner’s Dilemma, Intention-based Decision Making.

# 1 Introduction

*Intention recognition* can be found abundantly in many kinds of interactions and communications, not only in humans but also in many other species [88, 77, 40, 14]. It is so critical for human social functioning and the development of key human abilities, such as language and culture, that it is reasonable to assume it has been shaped by natural selection [88]. It is defined, in general terms, as the process of becoming aware of the intention of another agent and, more technically, as the problem of inferring an agent’s intention through its actions and their effects on the environment [34, 12, 26]. The knowledge about intention of others in a situation may enable to plan in advance, either to secure a successful cooperation or to deal with potential hostile behaviors [3, 83, 58], which may lead to dreadful collective outcomes [41, 61]. Intention recognition, and the recognition process or heuristics in general, are important mechanisms used by human bounded rationality to cope with real-life complexity [69, 70, 3, 66]. In contrast to the assumption of perfect rationality, we usually have to make decision in complicated and ill-defined situations, with incomplete information and time constraints [76, 70], under computational processing limitations, including limited memory capacity [70, 32, 31]. The recognition processes, based on stored knowledge in terms of generic patterns, have been known to play a major role in that respect. In problems with such complications, we look for patterns, based on them we simplify the problem by using these to construct temporary internal models as working hypotheses [3, 70]. It becomes even more relevant when considering interactive settings where the achievement of a goal by an agent does not depend solely on its own actions, but also on the decisions and actions of others, especially when the possibility of communication is limited [36, 26, 84]. The agents cannot rely on others to behave under perfect or improved rationality, and therefore need to be able to recognize their behaviors and even predict the intention beyond the surface behaviors. Additionally, in more realistic settings where deceiving may offer additional profits, individuals often attempt to hide their real intentions and make others believe in bogus ones [57, 77, 72, 52, 22, 63].

Hence, undoubtedly, a capability of intention recognition would confer on its holder great evolutionary benefits. However, although having been studied extensively in Artificial Intelligence, Philosophy and Psychology for several decades [34, 12, 9, 10, 51, 26, 20], few studies have addressed the evolutionary roles and aspects of intention recognition in regard to the emergence

of cooperation.

A crucial issue, traversing areas as diverse as Biology, Economics, Artificial Intelligence, Political Science, and Psychology, is to explain the *evolution of cooperation*: how natural selection can lead to cooperative behavior [23, 4, 68]. A cooperator may be seen as someone who pays a cost for another individual to receive a benefit. Cost and benefit are measured in terms of success or reproductive fitness. If two individuals simultaneously or independently decide to cooperate or not, the best possible response will be to try to receive the benefit without paying the cost. So, why does natural selection equip individuals with altruistic tendencies? Various mechanisms responsible for promoting such behavior have been recently proposed [68, 47], including kin and group ties [86, 78], different forms of reciprocity [44, 29, 82, 49, 46], and networked populations [60, 62, 73, 64]. In contradistinction, here we address explicitly how cooperation may emerge from the interplay between population dynamics and individuals’ cognitive abilities, namely the ability to perform intention recognition. As usual, our study is carried out within the framework of Evolutionary Game Theory (EGT) [27, 68].

In this work we take a first step towards employing intention recognition within an evolutionary setting and repeated interactions. As in the traditional framework of direct reciprocity [82], 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—given in terms of either a plan library [20, 12] or a plan corpus [7, 2, 21]. 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) for the different strategies. There have been several successful corpus-based intention recognition models in the literature [7, 8, 2], and we adjust a pre-existing one to the current work in lieu of supplying a completely novel one (see Subsection 2.5). The rationale of the corpus-based approach relies on the idea of nature-nurture co-evolution or experience inheritance [55, 67]: 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 with them—as suggested by the famous ‘like-me’ framework [39, 40]. 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 [28, 56, 42]. Closely related to this second point, notice that contrary to the formal EGT framework being used here, there have been several works showing the evolution of mind, which are based on genetic algorithm simulations, e.g. [75, 90, 65, 35]. Indeed, intention recognition can be considered as an elementary component of the “Theory of Mind” theory [53, 87].

In addition, to the best of our knowledge, although there is a large body of literature on learning in games [19, 37], e.g. reinforcement learning [59, 84], very little attention has been paid to studying how some cognitive ability that requires learning (as the ability of intention recognition in this work) fares in an evolutionary setting, particularly within the EGT framework. On the other hand, there have been some efforts to study the effects of increased memory sizes in evolutionary settings, e.g. see [24], though individual learning is not considered. Differently from this literature, our aim is to provide a computational model showing how the cognitive ability of intention recognition, which is so-critical and ubiquitous in humans’ activities [88, 77, 40], is a viable possibility that might have been retained by natural selection and co-evolves with the collective need to achieve high cooperative standards.

We offer a method to acquire an intention-based decision making model from the plan corpus, stating what to play with a given co-player based on the recognized intention and the game’s current state. The intention-based decision maker attempts to achieve the greatest expected benefit for himself/herself, taking advantage of the knowledge about the co-player’s intention. The model is discussed in Subsection 2.6.

In addition, we show that our intention recognizers prevail against the most famous successful strategies of repeated dilemmas of cooperation, including tit-for-tat, generous tit-for-tat and win-stay-lose-shift (described in Subsection 2.3), even in the presence of noise. Finally, as higher cognitive skills may imply a certain cost, we also discuss the implications of adding a (biological) cost of performing intention recognition.

## 2 Materials and Methods

### 2.1 Interaction between Agents

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

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \left( \begin{array}{cc} R, R & S, T \\ T, S & P, P \end{array} \right) \end{array}.$$

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 [37, 62, 68]. Namely, in this work we are concerned with the Prisoner’s Dilemma (PD), being characterized by  $T > R > P > S$ . In a single round, it is always best to defect, but cooperation may be rewarded if the game is repeated. In this case, it is additionally required that a mutual cooperation is preferred over an equal probability of unilateral cooperation and defection, i.e.  $2R > T + S$ ; otherwise alternating between cooperation and defection would lead to a higher payoff than mutual cooperation. For convenience and a clear representation of results, we later mostly use the Donor game [68]—a famous special case of the PD—where  $T = b$ ,  $R = b - c$ ,  $P = 0$ ,  $S = -c$ , satisfying that  $b > c > 0$ , where  $b$  and  $c$  stand respectively for “benefit” and “cost” (of cooperation).

We do not maintain that the repeated PD is necessarily a good model for the evolution of human cooperation [71, 6] but we nevertheless wanted to demonstrate the effectiveness of the intention recognition strategy against such well-explored PD strategies as tit-for-tat and win-stay-lose-shift (described in Subsection 2.3). In addition, because PD is the most difficult dilemma for cooperation to emerge, we envisage that the intention recognition strategy would do well in other dilemmas such as Stag-hunt game [71].

In a population of  $N$  individuals interacting via a repeated (or iterated) Prisoner’s Dilemma, whenever two specific strategies are present in the population, say **A** and **B**, the (per-round) payoff or fitness of an individual with a strategy **A** in a population with  $k$  **As** and  $(N - k)$  **Bs**

can be written as

$$\Pi_A(k) = \frac{1}{m(N-1)} \sum_{j=1}^m [(k-1)\pi_{A,A}(j) + (N-k)\pi_{A,B}(j)] \quad (1)$$

where  $\pi_{A,A}(j)$  ( $\pi_{A,B}(j)$ ) stands for the payoff obtained from a round  $j$  as a result of their mutual behavior of an **A** strategist in an interaction with a **A** (**B**) strategist (as specified by the payoff matrix above), and  $m$  is the total number of rounds of the Prisoner's Dilemma. As usual, instead of considering a fixed number of rounds, upon completion of each round, there is a probability  $w$  that yet another round of the game will take place, resulting in an average number of  $\langle m \rangle = (1-w)^{-1}$  rounds per interaction [68].

## 2.2 Evolutionary Dynamics

The accumulated payoff from all interactions (defined in Eq. (1)) emulates the individual *fitness* or social *success* and the most successful individuals will tend to be imitated by others, implementing a simple form of social learning [68]. A strategy update event is defined in the following way, corresponding to the so-called pairwise comparison [74, 79]. At each time-step, one individual  $i$  with a fitness  $f_i$  is randomly chosen for behavioral revision.  $i$  will adopt the strategy of a randomly chosen individual  $j$  with fitness  $f_j$  with a probability given by the Fermi function (from statistical physics)

$$p(f_i, f_j) = \left(1 + e^{-\beta(f_j - f_i)}\right)^{-1}$$

where the quantity  $\beta$ , which in physics corresponds to an inverse temperature, controls the intensity of selection. When  $\beta = 0$  we obtain the limit of neutral drift, and with the increasing of  $\beta$  one strengthens the role played by the game payoff in the individual fitness, and behavioral evolution [79, 80].

In the absence of mutations, the end states of evolution are inevitably monomorphic, as a result of the stochastic nature of the evolutionary dynamics and update rule. As we are interested in a global analysis of the population dynamics with multiple strategies, we further assume that with a small probability  $\mu$  individuals switch to a randomly chosen strategy, freely exploring the

space of possible behaviors. By introducing a small probability of mutation or exploration, the eventual appearance of a single mutant in a monomorphic population, this mutant will fixate or will become extinct long before the occurrence of another mutation and, for this reason, the population will spend all of its time with a maximum of two strategies present simultaneously [18, 29, 25, 63, 85]. This allows one to describe the evolutionary dynamics of our population in terms of a reduced Markov Chain of a size equal to the number of different strategies, where each state represents a possible monomorphic end-state of the population associated with a given strategy, and the transitions between states are defined by the fixation probabilities of a single mutant of one strategy in a population of individuals who adopt another strategy. The resulting stationary distribution characterizes the average time the population spends in each of these monomorphic states, and can be computed analytically [33, 18, 29, 25, 63, 85] (see below).

In the presence of two strategies the payoffs of each are given by Eq. (1), whereas the probability to change the number  $k$  of individuals with a strategy **A** (by  $\pm$  one in each time step) in a population of  $(N - k)$  **B**-strategists is

$$T^{\pm}(k) = \frac{N - k}{N} \frac{k}{N} \left[ 1 + e^{\mp \beta [\Pi_A(k) - \Pi_B(k)]} \right]^{-1}.$$

The fixation probability of a single mutant with a strategy **A** in a population of  $(N - 1)$  **B**s is given by [79]

$$\rho_{B,A} = \left( \sum_{i=0}^{N-1} \prod_{j=1}^i \lambda_j \right)^{-1} \quad (2)$$

where  $\lambda_j = T^-(j)/T^+(j)$ .

In the limit of neutral selection (i.e.,  $\beta = 0$ ),  $\lambda_j = 1$ . Thus,  $\rho_{B,A} = 1/N$ . Considering a set  $\{1, \dots, n_S\}$  of different strategies, these fixation probabilities determine a transition matrix  $[T_{ij}]_{i,j=1,\dots,n_S}$ , with  $T_{ii} = 1 - \sum_{k=1, k \neq i}^{n_S} \rho_{k,i}/(n_S - 1)$  and  $T_{ij, j \neq i} = \rho_{ji}/(n_S - 1)$ , of a Markov Chain. The normalized eigenvector associated with the eigenvalue 1 of the transposed of  $M$  provides the stationary distribution described above [33], describing the relative time the population spends adopting each of the strategies.



### 2.3 Strategies in Repeated Prisoner’s Dilemma

The repeated PD is usually known as a story of tit-for-tat (TFT), which won both Axelrod’s tournaments [4, 5]. *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 (e.g., when two *TFT*s playing with one another, a erroneous defection by one player leads to a sequence of unilateral cooperation and defection) 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  $p > 0$ ) [48]. *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 [45]. *WSLS* starts by cooperating, and 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 severely exploited by *AllD* (always defect) players.

### 2.4 Plan Corpus Description

In the following, we describe how to create plan corpora for training and testing the described intention recognition models, for a given set of strategies. For the sake of generality, we start by making 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  $M$ . To deal with such strategies, the corpus-based intention recognition strategy to be described is also equipped with a memory size of  $M$ , but no need for more. Note that the strategies described above (*TFT*, *GTFT*, *WSLS*) have memory bounded by  $M = 1$ , hence to cope with them, the intention recognizer memorizes a single last interaction with them.

For clarity of representation, abusing notations,  $R$ ,  $S$ ,  $T$  and  $P$  are henceforth also referred

to as (elementary) game states, in a single round of interaction. Additionally,  $E$  (standing for *empty*) is used to refer to a game state having had no interaction. The most basic elements in a plan corpus are the corpus actions, having the following representation.

**Definition 2.1 (Corpus Action)** *An action in a plan 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 <sup>1</sup>.*

**Definition 2.2 (Plan Session)** *A plan session of a strategy is a sequence of corpus actions played by that strategy (more precisely, a player using that strategy) against an arbitrary player.*

We denote by  $\Sigma_M$  the set of all possible types of action, such that  $|\Sigma_M| = 2 \times 5^M$ . For example, one can have

$$\Sigma_1 = \{EC, RC, TC, SC, PC, ED, RD, TD, SD, PD\}$$

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. The independence of actions will allow to derive a convenient and efficient intention recognition model, discussed in the next subsection. Furthermore, it enables to save the game states without having to save the co-player's moves, thus simplifying the representation of plan corpora.

As an example, let us consider *TFT* and the following sequence of its interactions with some other player (denoted by  $X$ ), in the presence of noise

$$\begin{array}{rcccccc} \text{round :} & 0 & 1 & 2 & 3 & 4 & 5 \\ \text{TFT :} & - & C & C & D & D & D \\ \text{X :} & - & C & D & D & C & D \\ \text{TFT-states :} & E & R & S & P & T & P \end{array}$$

The corresponding plan session for *TFT* is  $[EC, RC, SD, PD, TD]$ . At 0-th round, there is no interaction, thus the game state is  $E$ . *TFT* starts by cooperating (1-st round), hence the first

---

<sup>1</sup>From now on, this notion of an action is used, which is different of the notion of a move (C or D).

action of the plan session is EC. Since player  $X$  also cooperates in the 1-st round, the game state at this round is  $R$ .  $TFT$  reciprocates in the 2-nd round by cooperating, hence the second action of the plan session is  $RC$ . Similarly for the third and the fourth actions. Now, at the 5-th round,  $TFT$  should cooperate since  $X$  cooperated in 4-th round, but because of noise, it makes an error to defect. Therefore, the 5-th action is TD.

**Definition 2.3 (Plan Corpus)** *Let  $S$  be a set of strategies to be recognized. A plan corpus for  $S$  is a set of plan sessions generated for each strategy in the set.*

For a given set of strategies, different plan corpora can be generated for different purposes. In Subsection 3.1, for example, we generate plan corpora for training and testing intention recognition models.

## 2.5 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 described in [7, 8, 2]. In [7], the authors use the bigram statistical model, making the assumption that the current action only depends on the previous one. In [2], the authors attempt to avoid this assumption by using the Variable-Order Markov Model. In our work, because of the independence of (corpus) actions, we can derive an even simpler model than that in [7], as described below.

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

$$\begin{aligned} I^* &= \arg \max_{I_i: 1 \leq i \leq n} P(I_i | A_1 \dots A_m) \\ &= \arg \max_{I_i: 1 \leq i \leq n} \frac{P(I_i) \prod_{j=1}^m P(A_j | I_i, A_1 \dots A_{j-1})}{P(A_1 \dots A_m)} \end{aligned}$$

The second equation is obtained by applying the Bayes' and then Chain rules [50, 7]. Since the denominator  $P(A_1 \dots A_m)$  is a positive constant, we can ignore it. Then, because of the

independency amongst actions for the strategies being recognized, we obtain

$$I^* = \arg \max_{I_i: 1 \leq i \leq n} P(I_i) \prod_{j=1}^m P(A_j | I_i) \quad (3)$$

Note that this simplified expression is derived independently of the memory size  $M$ . All the probabilities needed for this computation are to be extracted beforehand using a training plan corpus. There is no update of these probabilities during intention recognizers' life cycle. These constants are arguably obtainable by evolutionary means via lineage [55, 67] – the corpus represents ancestors' experience in interacting with known strategies, or the 'like-me' self-experimenting framework [40, 39]. They are learnt once, and are used only for the purpose of performing intention recognition.

Also note that if two intentions are assessed with the same probability, the model predicts the one with higher priority. Priorities of intentions 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's intention can be predicted as either *AllC*, *WSLS* or *TFT* (since they are assigned the same conditional probability values). Being concerned of *TFT*'s and *WSLS*'s retaliation after a defect (i.e.  $IR$ 's behavior attitude), *WSLS* and *TFT* should have higher priorities than *AllC*.

## 2.6 Intention-based Decision Making Model

We describe how to acquire a decision making model for an intention recognizer from a training plan corpus. 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  $\theta(s, I)$ :

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

deciding what to play (C or D) given the sequence of  $M$  last game states  $s = s_1 \dots s_M$  ( $s_i \in \{E, R, T, S, P\}$  with  $1 \leq i \leq M$ ), and the recognized intention  $I \in \{I_1, \dots, I_n\}$ . It means that the intention recognizer needs to memorize the state of at most  $M$  last interactions with its co-player, besides the fixed set of probabilistic constants in Eq. (3) specifically for the purpose

of intention recognition. It is done as follows. From the plan sessions in the training corpus 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 sequence of states  $s$ . The choice giving greater payoff is chosen. Formally, let  $DS(I)$  be the set of all sequences of actions (plan sessions),  $Sq = A_1 \dots A_k$  ( $A_i \in \Sigma_M$ ,  $1 \leq i \leq k$ ), for intention  $I$  in the corpus and  $\pi(Sq, j)$  the payoff the intention recognizer would get at round  $j$ . In the following, if the sequence (plan session) in which the payoff being computed is clear from the context, we ignore it and simply write  $\pi(j)$ . Thus,

$$\theta(s, I) = \arg \max_{\xi \in \{C, D\}} \Pi_\xi / N_\xi \quad (4)$$

where

$$\Pi_\xi = \sum_{A_1 \dots A_k \in DS(I)} \sum_{\substack{1 \leq i \leq k \\ A_i = s\xi}} \sum_{j=i}^k \pi(j)$$

and

$$N_\xi = \sum_{A_1 \dots A_k \in DS(I)} \sum_{\substack{1 \leq i \leq k \\ A_i = s\xi}} \sum_{j=i}^k 1.$$

A more general version of this decision making model is provided by considering a discount factor, say  $1/\alpha$  where  $0 < \alpha \leq 1$ , stating how much the payoffs of distant rounds are less important:

$$\theta(s, I) = \arg \max_{\xi \in \{C, D\}} \Pi_{\alpha\xi} / N_{\alpha\xi} \quad (5)$$

where

$$\Pi_{\alpha\xi} = \sum_{A_1 \dots A_k \in DS(I)} \sum_{\substack{1 \leq i \leq k \\ A_i = s\xi}} \sum_{j=i}^k \alpha^{j-i} \pi(j)$$

and

$$N_{\alpha\xi} = \sum_{A_1 \dots A_k \in DS(I)} \sum_{\substack{1 \leq i \leq k \\ A_i = s\xi}} \sum_{j=i}^k \alpha^{j-i}.$$

The first model is a special case of the second one where  $\alpha = 1$ . This is the most future-optimistic case. If  $\alpha$  is small enough ( $\alpha \approx 0$ ),  $\theta(s, I) = D$ : in a given round of PD, it is always best to defect. This is the future-pessimistic case.

Note that, at the first round, there is no information about the co-player. The intention

recognizer cooperates, i.e.  $\theta(E^M, I) = C \quad \forall I \in \{I_1, \dots, I_n\}$ .

### 3 Models Acquisition and Evaluation

#### 3.1 Plan Corpus Generation

With the corpus description provided in Subsection 2.4, 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, this issue is beyond the scope of this article.

Let us further note that here we are not trying to design any optimal strategy for the repeated PD. Indeed, our aim is solely to provide a minimal model which supports the idea that intention recognition may prevail and evolve in the presence of cooperation dilemmas. While intention recognition may have played an important evolutionary role [88, 77, 39], it is difficult to conceive it as the key cognitive skill towards cooperation. Nevertheless, we could easily include other more complex strategies, such those with greater memory sizes and/or learning capability [38, 30], and let the intention recognizers learn to recognize them in a similar manner, by equipping the recognizers with memory capacity concomitant with that of its co-players’. As repeated games provide endless possibilities in what concerns the cognitive principles behind each particular strategy, as it can be easily understood, it was not our goal to provide an exhaustive study on all variants of the strategies analyzed here.

We collect plan sessions of each strategy by playing a random move (C or D) in each round with it. To be more thorough, we can also play all possible combinations for each given number of rounds  $m$ . For example, if  $m = 10$ , there will be  $1024 = 2^{10}$  combinations—C or D in each round. Moreover, since interactions in the presence of noise are taken into consideration, each combination is played repeatedly several times.

	<b>AllC</b>	<b>AllD</b>	<b>TFT</b>	<b>WSLS</b>	<b>Total</b>
<b>Precision</b>	0.859	<b>0.999</b>	0.818	0.579	<b>0.824</b>
<b>Recall</b>	0.859	0.999	0.818	0.575	0.824
<b>Converg.</b>	0.859	0.999	0.719	0.579	0.805

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

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

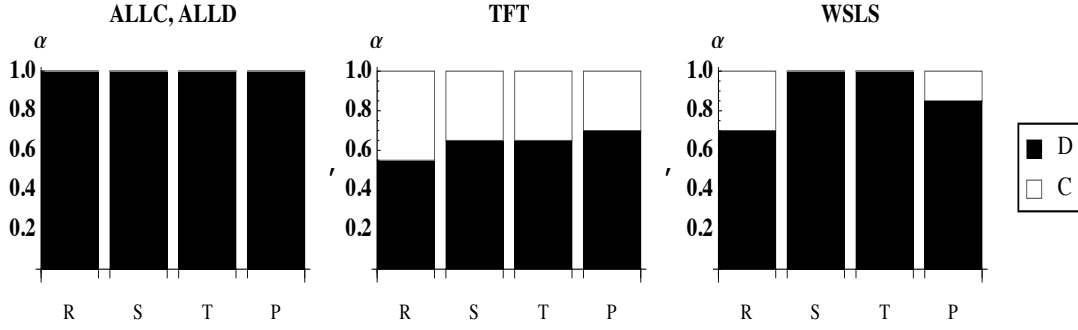
## 3.2 Intention Recognition Model

### 3.2.1 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 [2].

### 3.2.2 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 the 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 together with the intention-based decision making model (called *IR* players) in large scale population settings—particularly to address “What is the role of intention recognition for the emergence of cooperation?”



**Figure 1:** Decision making model for different values of  $\alpha$ . 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  $\alpha$  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  $\alpha$ . 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.

### 3.3 Decision Making Model

The decision making model (in Subsection 2.6) is acquired using the training corpus (Figure 1). We henceforth use  $\alpha = 1$ , i.e., independently of the PD payoff matrixes (used in this paper), if at the current round the co-player’s recognized intention is a unconditional one (*AllC* or *AllD*), IR always defects, regardless of the current game states; if it is *TFT*, IR always cooperates; and if it is *WSLS*, IR cooperates if and only if the current state is either *R* or *P*.

From here it is clear that, similar to *WSLS* and *TFT*, at each round, the strategy IR grounds its decision solely on the last move of others ( $M = 1$ ). The remaining memory capacity required for the functioning of IR is used for the purpose of identifying its co-player’s intention. That is to say, if IR is provided with its co-player’s intention, whether by any other possibly cheaper intention recognition methods or heuristics, it only needs to remember the last interaction. That means the success of our IR strategy, described below, is exclusively governed by the effect of taking into account its co-player’s intention in choosing a move, not its larger-memory capacity, used once to identify that intention<sup>2</sup>. In this respect, we also note the important role of intention-based decision making has been recognized in a diversity of experimental studies in behavioral economics, e.g. in [17, 16, 54].

<sup>2</sup>This should well be the case because the link between memory size and the most abstract aspects of intelligence must be an indirect one: conceptual thinking can be hardly reduced to simple memory storage and retrieval [15]. The development (and evolution) of complex cognitive abilities require long-term memory capacity, but the latter one is obviously not the whole story.



Figure 2 shows how an *IR* using these acquired intention recognition and intention-based decision making model interacts with other strategies, including *AllC*, *AllD*, *TFT*, *WSLS* and another *IR*, in the absence of noise. 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* is set with a higher priority than *WSLS*, which in turn 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* in the training corpus. 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. In playing with an *AllD*, *IR* cooperates in the first round, and defects in the remaining rounds, since  $P(ED|ALLD) \gg P(ED|I)$  for all  $I \in \{AllC, TFT, WSLS\}$  and furthermore  $P(s|ALLD) \gg P(s|I)$  for all  $I \in \{AllC, TFT, WSLS\}$  and  $s \in \{RD, SD, TD, PD\}$ .

## 4 Experiments and Results

In the following we provide analytical results under different evolutionary dynamics as well as using computer-based simulations. We show that the introduction of intention recognition promotes the emergence of cooperation in various settings, even in the presence of noise.

<b>IR:</b> C D D D D...	<b>IR:</b> C C D D D...	<b>IR:</b> C C D C C...
<b>AllD:</b> D D D D D...	<b>AllC:</b> C C C C C...	<b>IR:</b> C C D C C...
<b>IR:</b> C C D D C C C C...		<b>IR:</b> C C D D C C C C...
<b>TFT:</b> C C C D D C C C...	<b>WSLS:</b> C C C D C C C C...	

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

## 4.1 Analysis

To begin with, let us 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 the absence of noise, the payoff matrix of *AllC*, *AllD* and *IR* in  $m$  rounds is given by (Figure 2,  $\alpha = 1$ )

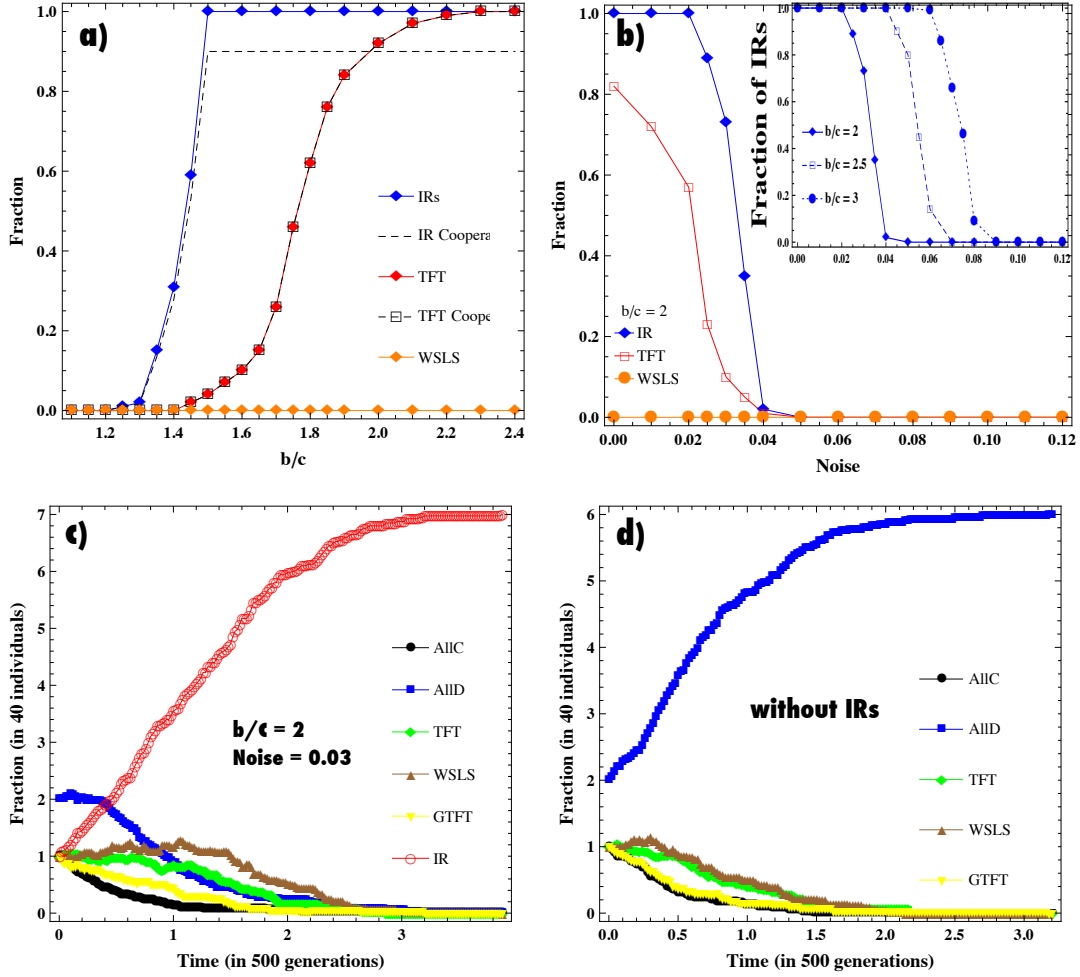
$$\begin{array}{c}
 \begin{array}{c} AllC \\ AllD \\ IR \end{array} \begin{pmatrix}
 \begin{array}{c} AllC \\ AllD \\ IR \end{array} \begin{array}{ccc}
 AllC & AllD & IR \\
 Rm & Sm & 2R + S(m-2) \\
 Tm & Pm & T + P(m-1) \\
 T(m-2) + 2R & P(m-1) + S & R(m-1) + P
 \end{array}
 \end{pmatrix}
 \end{array}$$

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 [27, 68, 47] of the three strategies lead to the conclusions that *AllC* is dominated by *IR* and that *IR* is an *evolutionary stable strategy* [27] if

$$R(m-1) > T + P(m-2) \quad (6)$$

which always holds for  $m \geq 3$  (since  $2R > T + P$  and  $R > P$ ). Evolutionarily stable strategy is a strategy which, if adopted by a population of players, cannot be invaded by any alternative strategy that is initially rare [27]. This condition guarantees that once *IR* dominates the population, it becomes stable (for  $m \geq 3$ ).



**Figure 3:** Simulation results for Donor game. In panels (a) and (b), we consider populations of three strategies, *AllC*, *AllD* and either *IR*, *TFT* or *WSLS*—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*, these two latter values 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 result shows that, in the presence of noise, *IR* outperforms *TFT* and *WSLS*. This result is robust to changes on the value of  $b/c$  (the inset of panel (b)). In panels (c) and (d), we consider a more complex setting where the population consists of several types of strategies: *AllC*, *AllD*, *TFT*, *WSLS*, *GTFT* (probability of forgiving a defect is 0.5) and *IR* (panel (c)) or without *IR* (panel (d)). Except for the defective *AllD* and *IR*, the other strategies are cooperative. Thus, instead of initially being equally distributed, we include a higher fraction of *AllDs* in the initial population. Namely, each type has 40 individuals, and *AllD* has 80. *IR* always wins (panel (c)). However, if *IR* individuals are removed, *AllD* is the winner (panel (d)), showing how *IRs* 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 *TFTs* as a catalyst to perform well [68]—which can be observed in panels (c) and (d). All results were obtained averaging over 100 runs, for  $m = 10$ ,  $N = 100$  and  $\beta = 0.1$ .

Furthermore, one can show that

1. *IR* is *risk-dominant* [27, 68] against *AllD* if  $R(m-1) + S > P(m-1) + T$ , which is equivalent to

$$m > \frac{T + R - S - P}{R - P}. \quad (7)$$

For Donor game, it is equivalent to:  $m > 2b/(b-c)$ .

2. *IR* is *advantageous* [27, 68] against *AllD* if  $R(m-1) + 2S > T + Pm$ , which is equivalent to

$$m > \frac{T + R - 2S}{R - P}. \quad (8)$$

For Donor game, it is equivalent to:  $m > (2b+c)/(b-c)$ .

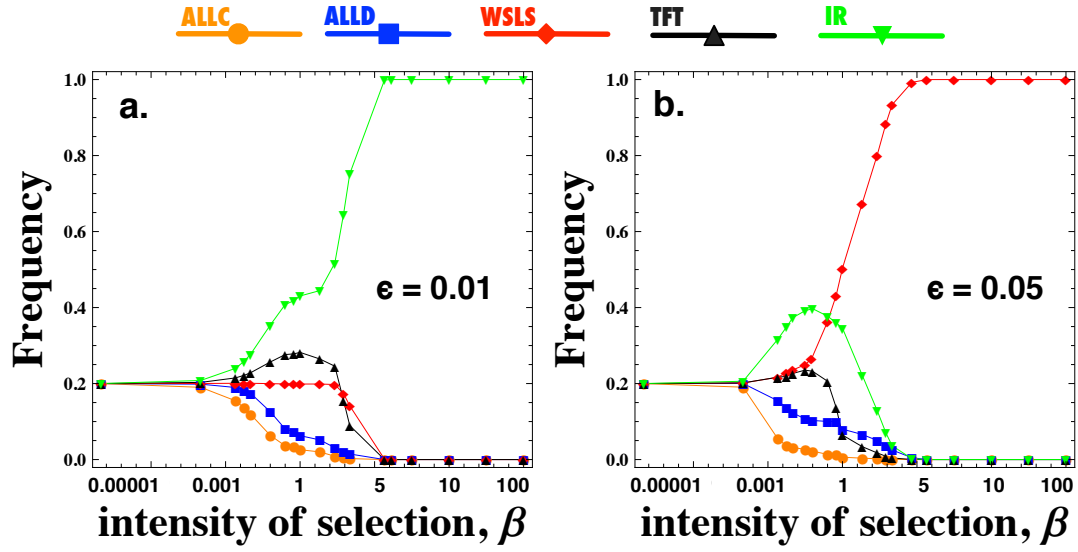
Since *IR* and *AllD* are both evolutionary stable strategies, Eq. (7) provides the condition for which *IR* has the greater basin of attraction [47, 68]. Eq. (8) provides the condition for which natural selection favors an *IR* to replace a population of *AllDs*, i.e. *IR* has a fixation probability greater than the neutral one ( $1/N$ ) [47, 68].

## 4.2 Evolutionary Simulations

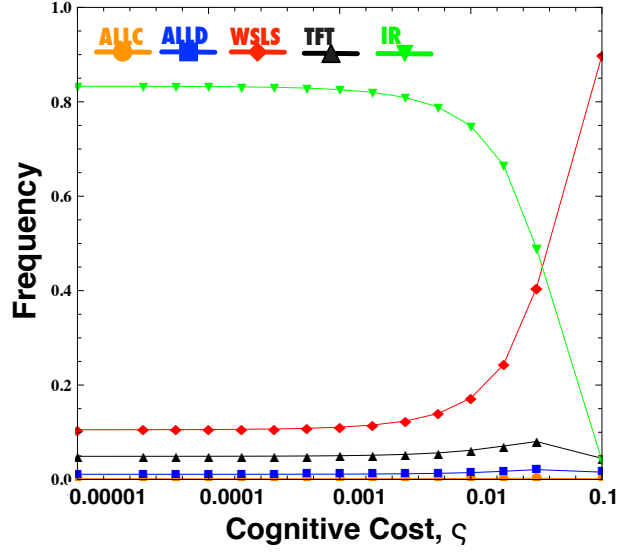
In the presence of noise, it was not easy to provide an exact mathematical analysis. Instead, we will study this case using computer simulations. For convenience and a clear representation of simulation results, we perform our simulations using the Donor game [68], i.e.,  $T = b$ ,  $R = b - c$ ,  $P = 0$ ,  $S = -c$ , satisfying that  $b > c > 0$ .

We start with a well-mixed population of size  $N$ , with individuals using different strategies. In each round of a generation, each individual interacts with all others, engaging in a PD game. The payoffs are accumulated over all the rounds. After each generation, an individual is randomly selected from the population, and will adopt the strategy of another randomly selected individual using the pairwise comparison rule [74, 79], described in Subsection 2.2.

The results for some different settings are shown in Figure 3. Our results show that *IR* always prevails against other strategies, including *TFT*, *WSLS* and *GTFT*, for different benefit-to-cost ratios  $b/c$ , as well as more robust to noise. Namely, it has a strictly larger range of the



**Figure 4:** Stationary distribution (in percentage) of each strategy depending on the intensity of selection  $\beta$ . The population consists of five strategies *ALLC*, *ALLD*, *TFT*, *WSLS*, and *IR*. For small values of  $\beta$ , selection is nearly neutral. Strategy updating is mostly random and frequencies of all strategies are roughly equal. Discrimination between strategies occurs when  $\beta$  increases. (a) When noise is small, *IR* always wins; (b) When noise is large, *IR*s wins for small  $\beta$ ; *WSLS* wins when  $\beta$  is large. All calculations are made with  $b/c = 3$ ,  $m = 10$ ,  $N = 100$ . When noise is present, the average payoffs of each strategy are obtained by averaging  $10^7$  runs.



**Figure 5:** Stationary distribution (in percentage) of each strategy depending on the cognition cost,  $\varsigma$ . The population consists of five strategies *AllC*, *AllD*, *TFT*, *WSLS*, and *IR*. For small enough values of  $\varsigma$ , the population spends most of the time in the homogenous state of *IR*. When  $\varsigma$  is large, *WSLS* prevails. All calculations are made with  $b = 3$ ,  $c = 1$ ,  $m = 10$ ,  $N = 100$ ,  $\epsilon = 0.01$ . When noise is present, the average payoffs of each strategy are obtained by averaging  $10^7$  runs.

benefit-to-cost ratio where cooperation can emerge, and can maintain it under a larger level of noise.

### 4.3 Intensity of Selection

We consider a setting where five strategies *AllC*, *AllD*, *TFT*, *WSLS*, and *IR* are present in the population. We compute numerically stationary distributions for variable intensity of selection  $\beta$  (Figure 4), using the method described in Subsection 2.2. The results show that, for small enough noise, the population always spends more time in the homogeneous state of *IR*, especially for strong intensities of selection (Figure 4a). When noise is large, *WSLS* wins for strong intensities of selection, but *IR* still wins for the slow ones (Figure 4b).

Note that in case the intensity of selection  $\beta$  is very small, it is commonly referred to in the literature as *weak selection* [43, 79, 89]. Weak selection describes the situation in which the returns from the game represent a small perturbation to the fitness of an individual [1, 11, 81].

#### 4.4 Cognitive Cost of Intention Recognition

We now extend our model to take into account the (cognitive) cost needed to perform intention recognition. Let us denote this cost of cognition by  $\varsigma$ . In each interaction, it is subtracted from the payoff of *IRs*. Given the difficulty to assess the costs associated with a given cognitive skill—see, e.g., [38, 30, 13], where cognition costs were not considered as significant in the time-scale of strategy evolution—we are not able to provide a particular value, even if in relative terms. For this reason, we provide analysis on the range of  $\varsigma$  under which *IRs* still prevail, identifying the values of  $\varsigma$  above which *IR* will, as expected, be washed out from the population.

In a population of three strategies, *AllC*, *AllD* and *IR*, and in the absence of noise, we derive some analytical results for  $\varsigma$  under which *IR* is favored by natural selection in different settings. First, we can re-write the payoff matrix as follows

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

Similarly to the analysis in Section 4.1, pairwise comparisons of the three strategies lead to the conclusions that *IR* is an *evolutionary stable strategy* if

$$\varsigma < \min \left\{ R - S - \frac{3R - P - 2S}{m}, R - P - \frac{R + T - 2P}{m} \right\}. \quad (9)$$

This inequality holds if  $m \geq 3$  (guaranteeing that the right hand side is positive) and the cognitive cost  $\varsigma$  is small enough. These conditions also guarantee that *AllC* is dominated by *IR*.

One can show that *IR* is *risk-dominant* [27, 68] against *AllD* if

$$\varsigma < \frac{R - P}{2} - \frac{T + R - S - P}{2m}. \quad (10)$$

Furthermore, *IR* is *advantageous* [27, 68] against *AllD* if

$$\varsigma < \frac{R - P}{3} - \frac{T + R - 2S}{3m}. \quad (11)$$

Eq. (10) provides the condition on  $\varsigma$  for which *IR* has the greater basin of attraction [47, 68]. Furthermore, Eq. (11) provides the condition on  $\varsigma$  for which natural selection favors an *IR* to replace a population of *AllDs*, i.e. *IR* has a fixation probability greater than the neutral one ( $1/N$ ) [47, 68].

Now let us consider the population where all five strategies *AllC*, *AllD*, *TFT*, *WSLS*, and *IR* are present in the population. Our simulation result (Figure 5) shows that if the cost to perform intention recognition is small enough, *IR* prevails: the population spends most of the time in the homogeneous state *IR*. However, as  $\varsigma$  increases, as expected, the advantage of *IR* is undermined by this external cost, letting however the conditions for cooperation to prevail in the long-run, as *IR* is replaced by *WSLS*.

## 5 Concluding Remarks

Using the tool of evolutionary game theory, we have shown, analytically as well as by extensive computer simulations, that intention recognition may co-evolve with cooperation. From this co-evolution, cooperation prevails in the long-run. 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. Individuals which are equipped with an ability to recognize intention of others, i.e. intention recognizers, can quickly recognize the unconditional defectors *AllDs*, thus not being exploited by them as for the *WSLS* strategy. For their own benefit, the intention recognizers can recognize and exploit the unconditional cooperators *AllCs*, thus do not suffer random drift as *TFT* and *GTFT*. Furthermore, the intention recognizers are cooperative with the conditional cooperators, including *TFT*, *WSLS*, and players like themselves.

We have shown that a population with some initial fraction of intention recognizers acting selfishly to achieve greatest benefit can lead to a stable cooperation where intention recognizers come to prevail upon and permeate the population. The intention recognition strategy has a greater range of benefit-to-cost ratios leading to cooperation than the most successful existent strategies, including *TFT* and *WSLS*. We have also shown that it is more robust to noise, and it prevails under a various range of intensities of selection. Furthermore, all mentioned results are robust assuming that the cost of performing intention recognition is small enough compared



to the cost of attempting to cooperate. This outcome is only undermined in the presence of external factors, represented here by a cost associated with higher cognitive skills.

In addition, our approach of using a plan corpus makes a case for other artificial intelligence 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, as pattern recognition algorithms, and others. 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). Bridging such sophisticated techniques with human evolution and behavior remains an open challenge, both for artificial intelligence and theoretical biology. Following an already long tradition in artificial life research, our study provides a step further in that direction.

## Acknowledgments

TAH and FCS acknowledge the support from FCT-Portugal (grant SFRH/BD/62373/2009 and R&D project PTDC/FIS/101248/2008, respectively).

## References

- [1] H. Akashi. Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in *Drosophila* DNA. *Genetics*, 139(2):1067–1076, 1995.
- [2] Marcelo G. Armentano and Analía Amandi. Goal recognition with variable-order markov models. In *Proceedings of the 21st international joint conference on Artificial intelligence*, pages 1635–1640. Morgan Kaufmann, 2009.
- [3] W. B. Arthur. Inductive reasoning and bounded rationality. *American Economic Review*, 84(2):406–411, 1994.
- [4] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, ISBN 0-465-02122-2, 1984.
- [5] Robert Axelrod and W.D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.
- [6] K.G. Binmore. *Natural justice*. Oxford University Press, 2005.
- [7] N. Blaylock and J. Allen. Corpus-based, statistical goal recognition. In *Proceedings of the 18th international joint conference on Artificial intelligence (IJCAI'03)*, pages 1303–1308, 2003.
- [8] N. Blaylock and J. Allen. Statistical goal parameter recognition. In Shlomo Zilberstein, Jana Koehler, and Sven Koenig, editors, *Proceedings of the 14th International Conference on Automated Planning and Scheduling (ICAPS04)*, pages 297–304. AAAI, 2004.
- [9] M. E. Bratman. *Intention, Plans, and Practical Reason*. The David Hume Series, CSLI, 1987.
- [10] M. E. Bratman. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press, 1999.
- [11] Jane Charlesworth and Adam Eyre-Walker. The other side of the nearly neutral theory, evidence of slightly advantageous back-mutations. *Proc. Natl. Acad. Sci. USA*, 104(43):16992–16997, 2007.

- [12] E. Charniak and R.P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, 64(1):53–79, 1993.
- [13] A.W. Delton, M.M. Krasnow, L. Cosmides, and J. Tooby. Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proc Natl Acad Sci U S A*, 108(32):13335–40, 2011.
- [14] R. Dunbar. *Science of Love and Betrayal*. Faber & Faber, Limited, 2012.
- [15] Joël Fagot and Robert G. Cook. Evidence for large long-term memory capacities in baboons and pigeons and its implications for learning and the evolution of cognition. *Proceedings of the National Academy of Sciences*, 103(46):17564–17567, 2006.
- [16] Armin Falk, Ernst Fehr, and Urs Fischbacher. Testing theories of fairness—intentions matter. *Games and Economic Behavior*, 62(1):287 – 303, 2008.
- [17] Robert H. Frank, Thomas Gilovich, and Dennis T. Regan. The evolution of one-shot cooperation: An experiment. *Ethology and Sociobiology*, 14(4):247 – 256, 1993.
- [18] D. Fudenberg and L. A. Imhof. Imitation processes with small mutations. *Journal of Economic Theory*, 131:251–262, 2005.
- [19] Drew Fudenberg and David K. Levine. *The Theory of Learning in Games*, volume 1 of *MIT Press Books*. The MIT Press, December 1998.
- [20] Christopher W. Geib and Robert P. Goldman. A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence*, 173(2009):1101–1132, 2009.
- [21] T. A. Han and L. M. Pereira. Context-dependent incremental intention recognition through Bayesian network model construction. In A. Nicholson, editor, *Proceedings of the Eighth UAI Bayesian Modeling Applications Workshop (UAI-AW 2011)*, volume 818, pages 50–58. CEUR Workshop Proceedings, 2011.
- [22] T. A. Han and L. M. Pereira. Intention-based decision making with evolution prospection. In L. Antunes and H.S. Pinto, editors, *Proceedings of 15th Portuguese Conference on Artificial Intelligence (EPIA’2011)*, pages 254–267. Springer LNAI 7026, 2011.

- [23] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.
- [24] C. Hauert and H. Schuster. Effects of increasing the number of players and memory size in the iterated prisoner’s dilemma: a numerical approach. *Royal Society B: Biological Sciences*, 264(1381):513–519, 1997.
- [25] C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund. Via freedom to coercion: The emergence of costly punishment. *Science*, 316:1905–1907, 2007.
- [26] C. Heinze. *Modeling Intention Recognition for Intelligent Agent Systems*. PhD thesis, The University of Melbourne, Australia, 2003.
- [27] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.
- [28] M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J. C. Mazziotta, and G. Rizzolatti. *PLoS Biol*, 3(3):e79, 2005.
- [29] L. A. Imhof, D. Fudenberg, and Martin A. Nowak. Evolutionary cycles of cooperation and defection. *Proc Natl Acad Sci USA*, 102:10797–10800, 2005.
- [30] M Janssen. Evolution of cooperation in a one-shot prisoner’s dilemma based on recognition of trustworthy and untrustworthy agents. *Journal of Economic Behavior and Organization*, 65(3-4):458–471, 2008.
- [31] Philip N. Johnson-Laird. Mental models and human reasoning. *Proc. Natl. Acad. Sci. USA*, 107(43):18243–18250, 2010.
- [32] Yaakov Kareev. Through a narrow window: working memory capacity and the detection of covariation. *Cognition*, 56(3):263 – 269, 1995.
- [33] S. Karlin and H. E. Taylor. *A First Course in Stochastic Processes*. Academic Press, 1975.
- [34] H. Kautz and J.F. Allen. Generalized plan recognition. In *Proceedings of the Conference of the American Association of Artificial Intelligence (AAAI-86)*, pages 32–38. AAAI, 1986.

- [35] Kyung-Joong Kim and Hod Lipson. Towards a "theory of mind" in simulated robots. In *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, GECCO '09, pages 2071–2076, New York, NY, USA, 2009. ACM.
- [36] S. Kraus. Negotiation and cooperation in multi-agent environments. *Artificial Intelligence*, 94(1-2):79–98, 1997.
- [37] M. W. Macy and A. Flache. Learning dynamics in social dilemmas. *Proc. Natl. Acad. Sci. USA*, 99:7229–7236, 2002.
- [38] Luke McNally, Sam P. Brown, and Andrew L. Jackson. Cooperation and the evolution of intelligence. *Proceedings of the Royal Society B: Biological Sciences*, 279(1740):3027–3034, 2012.
- [39] A. N. Meltzoff. Imitation and other minds: the "like me" hypothesis. In *Perspectives on imitation: From neuroscience to social science. Imitation, human development, and culture*, pages 55–77. Cambridge, MA: MIT Press, 2005.
- [40] A. N. Meltzoff. The framework for recognizing and becoming an intentional agent. *Acta Psychol Amst*, 124(1):26–43, 2007.
- [41] M. Milinski, D. Semmann, H. Krambeck, and J. Marotzke. Stabilizing the earths climate is not a losing game: Supporting evidence from public goods experiments. *Proc. Natl. Acad. Sci. USA*, 103 (11):3994–3998, 2006.
- [42] Kiyoshi Nakahara and Yasushi Miyashita. Understanding intentions: Through the looking glass. *Science*, 308(5722):644–645, 2005.
- [43] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428:646–650, 2004.
- [44] M. A. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355:250–253, 1992.

- [45] M. 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.
- [46] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437(1291-1298), 2005.
- [47] Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560, 2006. DOI: 10.1126/science.1133755.
- [48] Martin A. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355:250–253, 1992.
- [49] J. M. Pacheco, F. C. Santos, and F. A. C. C. Chalub. Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comput. Biol.*, 2(12):e178, 2006.
- [50] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.
- [51] L. M. Pereira and T. A. Han. Intention recognition via causal bayes networks plus plan generation. In *Progress in Artificial Intelligence, Proceedings of 14th Portuguese International Conference on Artificial Intelligence (EPIA ’09)*, pages 138–149. Springer LNAI 5816, October 2009.
- [52] Luís Moniz Pereira and The Anh Han. Intention recognition with evolution propection and causal bayesian networks. In Ana Madureira, Judite Ferreira, and Zita Vale, editors, *Computational Intelligence for Engineering Systems: Emergent Applications*, volume 46, pages 1–33. Springer, 2011.
- [53] D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4):515–526, 1978.
- [54] Sina Radke, Berna Guroglu, and Ellen R. A. de Bruijn. There’s something about a fair split: Intentionality moderates context-based fairness considerations in social decision-making. *PLoS ONE*, 7(2):e31491, 02 2012.

- [55] Peter J. Richerson and Robert Boyd. *Not By Genes Alone: How Culture Transforms Human Evolution*. The University of Chicago Press, 2006.
- [56] Giacomo Rizzolatti and Laila Craighero. The mirror-neuron system. *Annual Review of Neuroscience*, 27:169–192, 2004.
- [57] A. Robson. Efficiency in evolutionary games: Darwin, nash, and the secret handshake. *Journal of Theoretical Biology*, 144(3):379–396, 1990.
- [58] Olivier Roy. Intentions and interactive transformations of decision problems. *Synthese*, 169(2):335–349, 2009.
- [59] Tuomas W. Sandholm and Robert H. Crites. Multiagent reinforcement learning in the iterated prisoner’s dilemma. *Biosystems*, 37:147–166, 1995.
- [60] F. C. Santos and J. M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.*, 95:098104, 2005.
- [61] F. C. Santos and J. M. Pacheco. Risk of collective failure provides an escape from the tragedy of the commons. *Proc. Natl. Acad. Sci. USA*, 108:doi: 10.1073.pnas.1015648108, 2011.
- [62] F. C. Santos, J. M. Pacheco, and T. Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. U.S.A.*, 103:3490–3494, 2006.
- [63] F. C. Santos, J. M. Pacheco, and B. Skyrms. Co-evolution of pre-play signaling and cooperation. *Journal of Theoretical Biology*, 274(1):30–35, 2011.
- [64] F. C. Santos, M. D. Santos, and J. M. Pacheco. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 454:214–216, 2008.
- [65] Hiroki Sayama, Dene L. Farrell, and Shelley D. Dionne. The effects of mental model formation on group decision making: An agent-based simulation. *Complexity*, 16(3):49–57, 2011.
- [66] R. Selten. What is bounded rationality? In *Bounded rationality: The adaptive toolbox*, pages 13–36. Cambridge, MA: MIT Press, 2001.

- [67] Stephen Shennan. *Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution*. London: Thames and Hudson Ltd., 2002.
- [68] Karl Sigmund. *The Calculus of Selfishness*. Princeton U. Press, 2010.
- [69] H. A. Simon. A behavioral model of rational choice. In *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley, 1957.
- [70] H. A. Simon. Invariants of human behavior. *Annual review of psychology*, 41:1–19, 1990.
- [71] Brian Skyrms. *The stag hunt and the evolution of social structure*. Cambridge University Press, 2004.
- [72] Brian Skyrms. *Signals: Evolution, Learning, and Information*. Oxford University Press, 2010.
- [73] G. Szabó and G. Fáth. Evolutionary games on graphs. *Phys Rep*, 97-216(4-6), 2007.
- [74] G. Szabo and C. Toke. Evolutionary prisoners dilemma game on a square lattice. *Phys Rev E*, 58:69–73, 1998.
- [75] M. Takano and T. Arita. Asymmetry between Even and Odd Levels of Recursion in a Theory of Mind. pages 405–411. Proceedings of ALife X, 2006.
- [76] P.M. Todd. Fast and frugal heuristics for environmentally bounded minds. In *Bounded rationality: The adaptive toolbox*, pages 51–70. Cambridge, MA: MIT Press, 2001.
- [77] M. Tomasello. *Origins of Human Communication*. MIT Press, 2008.
- [78] A. Traulsen and M. A. Nowak. Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. U.S.A.*, 103(29):10952, 2006.
- [79] A. Traulsen, M. A. Nowak, and J. M. Pacheco. Stochastic dynamics of invasion and fixation. *Phys. Rev. E*, 74:11909, 2006.
- [80] A. Traulsen, J. M. Pacheco, and M. A. Nowak. Pairwise comparison and selection temperature in evolutionary game dynamics. *J. Theor. Biol.*, 246:522–529, 2007.



- [81] A. Traulsen, D. Semmann, R.D. Sommerfel, H. Krambeck, and M. Milinski. Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci. U.S.A.*, 107(7):2962–2966, 2010.
- [82] R. L. Trivers. The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46:35–57, 1971.
- [83] M. van Hees and O. Roy. Intentions and plans in decision and game theory. In *Reasons and intentions*, pages 207–226. Ashgate Publishers, 2008.
- [84] S. Van Segbroeck, S. de Jong, A. Nowe, F. C. Santos, and T. Lenaerts. Learning to coordinate in complex networks. *Adaptive Behavior*, 18:416–427, 2010.
- [85] S. Van Segbroeck, J. M. Pacheco, T. Lenaerts, and F. C. Santos. Emergence of fairness in repeated group interactions. *Phys. Rev. Lett.*, 108:158104, 2012.
- [86] S.A. West, A.A. Griffin, and A. Gardner. Evolutionary explanations for cooperation. *Current Biology*, 17:R661–R672, 2007.
- [87] A. Whiten. *Natural theories of mind: evolution, development, and simulation of everyday mindreading*. B. Blackwell, 1991.
- [88] Amanda L. Woodward, Jessica A. Sommerville, Sarah Gerson, Annette M.E. Henderson, and Jennifer Buresh. The emergence of intention attribution in infancy. In Brian H. Ross, editor, *The Psychology of Learning and Motivation*, volume 51 of *Psychology of Learning and Motivation*, pages 187 – 222. Academic Press, 2009.
- [89] Bin Wu, Philipp M. Altrock, Long Wang, and Arne Traulsen. Universality of weak selection. *Phys. Rev. E*, 82(4):046106, Oct 2010.
- [90] F. Zanlungo. *Microscopic dynamics of artificial life systems*. PhD thesis, 2007.

	<b>AII</b>	<b>AII</b>	<b>TFT</b>	<b>WSLS</b>	<b>Total</b>
<b>Precision</b>	0.859	<b>0.999</b>	0.818	0.579	<b>0.824</b>
<b>Recall</b>	0.859	0.999	0.818	0.575	0.824
<b>Converg.</b>	0.859	0.999	0.719	0.579	0.805

**Table 1:**

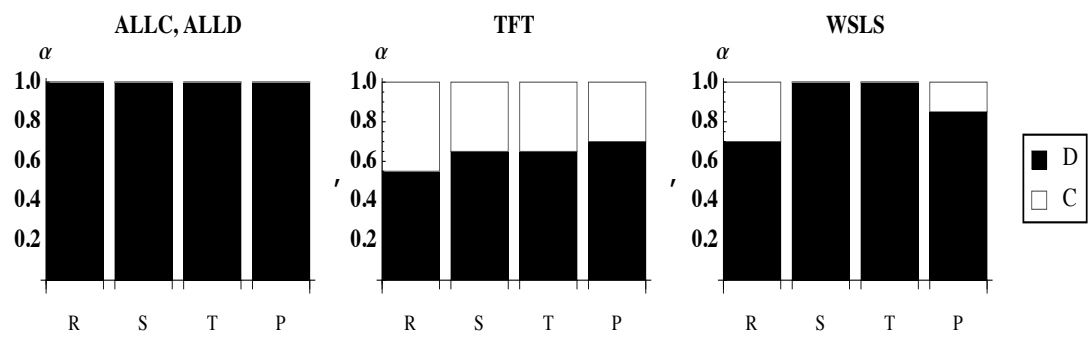


Figure 1:

<b>IR:</b> C D D D D...	<b>IR:</b> C C D D D...	<b>IR:</b> C C D C C...
<b>AIID:</b> D D D D D...	<b>AIIC:</b> C C C C C...	<b>IR:</b> C C D C C...
<b>IR:</b> C C D D C C C C...	<b>IR:</b> C C D D C C C C...	
<b>TFT:</b> C C C D D C C C...	<b>WSLS:</b> C C C D C C C C...	

**Figure 2:**

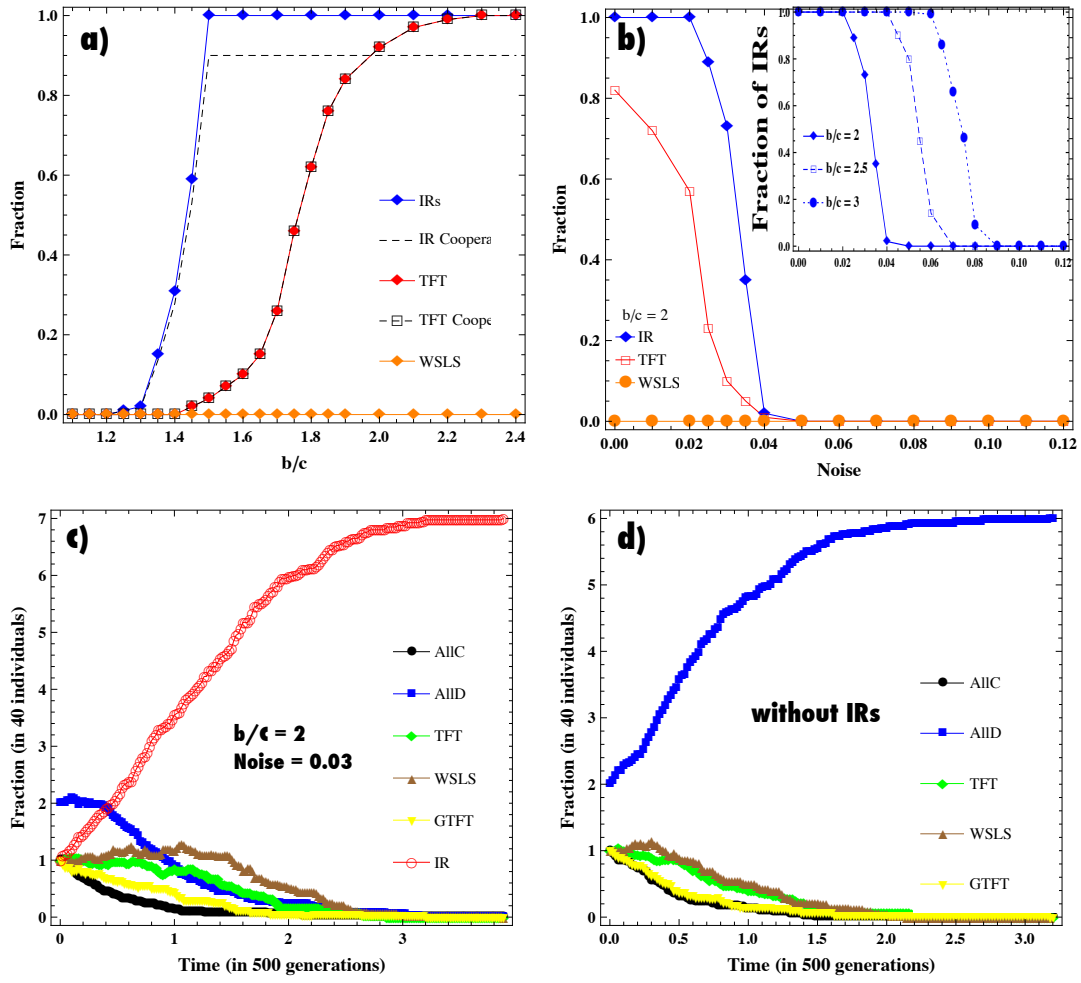


Figure 3

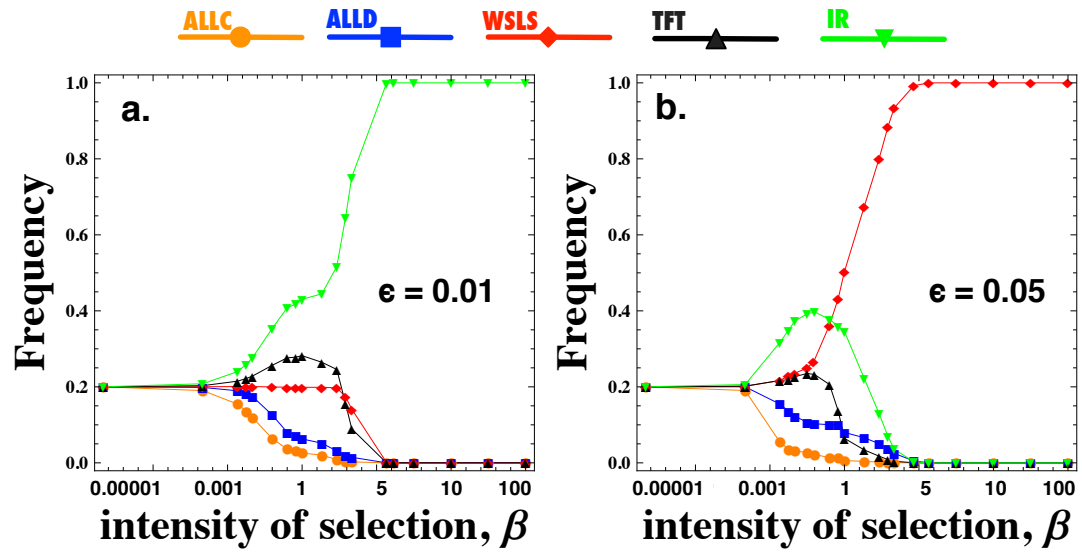


Figure 4:

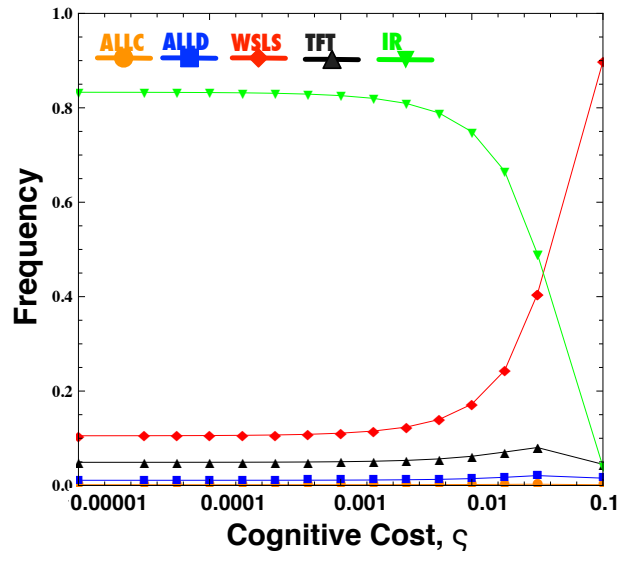


Figure 5: