The Knowledge Engineering Review, Vol. 23:2, 153–180. © 2008, Cambridge University Press
doi:10.1017/S026988890800132X Printed in the United Kingdom

Fairness in multi-agent systems

STEVEN DE JONG[1], KARL TUYLS[1,2] and KATJA VERBEECK[3]

[1]*MICC, Maastricht University, The Netherlands;*
e-mail: steven.dejong@micc.unimaas.nl;
[2]*Eindhoven Technical University, The Netherlands;*
e-mail: k.p.tuyls@tue.nl;
[3]*Katholieke Hogeschool Sint Lieven, Gent, Belgium;*
e-mail: katja.verbeeck@kahosl.be

Abstract

Multi-agent systems are complex systems in which multiple autonomous entities, called agents, cooperate in order to achieve a common or personal goal. These entities may be computer software, robots, and also humans. In fact, many multi-agent systems are intended to operate in cooperation with or as a service for humans. Typically, multi-agent systems are designed assuming perfectly rational, self-interested agents, according to the principles of classical game theory. Recently, such strong assumptions have been relaxed in various ways. One such way is explicitly including principles derived from human behavior. For instance, research in the field of behavioral economics shows that humans are not purely self-interested. In addition, they strongly care about *fairness*. Therefore, multi-agent systems that fail to take fairness into account, may not be sufficiently aligned with human expectations and may not reach intended goals. In this paper, we present an overview of work in the area of fairness in multi-agent systems. More precisely, we first look at the classical agent model, that is, rational decision making. We then provide an outline of descriptive models of fairness, that is, models that explain how and why humans reach fair decisions. Then, we look at prescriptive, computational models for achieving fairness in adaptive multi-agent systems. We show that results obtained by these models are compatible with experimental and analytical results obtained in the field of behavioral economics.

1 Introduction

In our increasingly interconnected world, research in the area of multi-agent systems is becoming more and more important (Jennings *et al*., 1998; Ferber, 1999; Weiss, 1999; Shoham *et al*., 2007). Multi-agent systems are generally accepted as valuable tools for designing and building distributed dynamical systems, by using several interacting agents, possibly including humans.

While there is no general definition for concepts such as agents and multi-agent systems, we adopt the following definitions: (1) an *agent* is an entity, situated in some environment, that is capable of flexible autonomous action in order to meet its objectives; (2) a *multi-agent system* is 'a loosely coupled network of agents that work together to solve problems that are beyond the individual capabilities or knowledge of each agent' (Jennings *et al*., 1998). Note that we may consider humans to be agents as well. In practice, multi-agent systems are often performing tasks in cooperation with, or instead of humans. Examples include software agents participating in online auctions or bargaining (Erev and Roth, 1998; Preist and van Tol, 1998; Kalagnanam and Parkes, 2004), electronic institutions (Rodriguez-Aguilar, 2003), developing schedules for air traffic (Mao *et al*., 2006), or decentralized resource distribution in large storage facilities (Weyns *et al*., 2005; De Jong *et al*., 2006).

Although multi-agent systems have many potential advantages, designing them raises many difficulties. One of the key problems lies in controlling the behavior of individual agents in such

a way that the system as a whole reaches a certain goal. This problem becomes even more prominent in multi-agent systems that interact with humans. Typically, multi-agent systems are designed assuming perfectly rational, self-interested agents, according to the principles of classical game theory. However, recently, these strong assumptions have been relaxed in various ways, for instance by including well-known concepts such as bounded rationality (Simon, 1957, 1972; Simon and Newell, 1972) and social welfare (Chevaleyre *et al.*, 2006, 2007). We feel that multi-agent-systems research may benefit even more by explicitly including principles derived from human behavior. Research in the field of behavioral economics shows us that humans are not purely rational and self-interested; their decisions are often based on considerations about others. For instance, humans strongly care about receiving a fair share (Bowles *et al.*, 1997; Fehr and Schmidt, 1997; Gintis, 2001). Therefore, multi-agent systems that fail to take such considerations into account, may not be sufficiently aligned with the humans they are interacting with. As a result, these systems might not reach their intended goals.

This phenomenon has been analyzed extensively from a game-theoretic context, using games such as the well-known Ultimatum Game (Fehr and Schmidt, 1999), Public Goods Game (Dannenberg *et al.*, 2007) and Traveler's Dilemma (Basu, 1994, 2007) in which a perfectly rational player is willing to accept a very low payoff and expects others to do the same, which will lead to a payoff that is much lower than the payoff typically obtained by human players.[1] For instance, in the Traveler's Dilemma, two players (travelers) independently have to report on the value of lost luggage. The lowest reported value is subsequently paid to the players, with an additional payoff of \$2 to the player who reported this value, and a penalty of \$2 to the other player. As a result, if a player assumes that the other player will report a value of x, he obtains the best payoff ($x+1$) by reporting a value of $x-1$. Thus, the other player is better off by reporting $x-2$ (and obtaining x instead of $x-2$). Continuing along this line of reasoning, a perfectly rational player will report a value of \$2. Humans consistently report values close to the maximum allowed, e.g., \$98 in case of the maximum being \$100. Thus, typical human players obtain a payoff of \$90 or more, whereas purely rational players obtain only \$2. More generally speaking, being aware of fairness may lead to better results in any problem domain in which the allocation of limited resources plays an important role (Chevaleyre *et al.*, 2006), as in many of the aforementioned examples.

Thus, agents in many multi-agents systems should be made explicitly aware of potentially irrational concepts that humans care about. In practice, this means that concepts such as fairness, discovered in such diverse fields as behavioral economics, economical psychology and evolutionary game theory, must be well-understood by developers of multi-agent systems. Fairness has been extensively studied, resulting in so-called *descriptive* models of human fairness, explaining why and how humans reach fair solutions instead of individually rational ones. These models may be used as a basis for *prescriptive* or computational models, used to control agents in multi-agent systems in such a way that alignment with human expectations is achieved. In this paper, we aim at providing an overview of work in this interesting and interdisciplinary area. First, we give a concise overview of the 'classical' agent model, i.e., rational decision making and game theory, introducing concepts such as Nash equilibria and Pareto-optimality. Second, we discuss three descriptive fairness models, being inequity aversion, priority awareness and reciprocal fairness. Third, we discuss computational fairness models, i.e., models that can be used to make agents in multi-agent systems reach fair, 'human' solutions. Finally, we conclude.

2 Rational decision-making

In this section, we give a brief overview of the 'classical' agent model, i.e., rational decision making and game theory. For a detailed overview, we refer to (Osborne and Rubinstein, 1994; Tuyls and Nowe, 2005).

[1] The intuitions behind this for the Ultimatum and Public Goods Game will be explained in Section 3.1.2.

Games, actions and utility. Game theory is used to model and analyze various strategic interactions between different people (or agents) (von Neumann and Morgenstern, 1944). Interactions are modeled in the form of games. More precisely, we assume a collection of n agents where each agent i has an individual finite set of actions A_i. The number of actions in A_i is denoted by $|A_i|$. In many real-world scenarios, for instance in bargaining, agents actually have a very large (or possibly infinite) set of actions. Traditionally however, game theory assumes a discrete, finite set of actions available to every agent.

The agents play a game in which each agent i independently selects an individual action a from its private action set A_i. The combination of actions of all agents constitute a joint action or action profile $\vec{a}$ from the joint action set $\mathbb{A} = A_1 \times \ldots \times A_n$. A joint action $\vec{a}$ is thus a vector in the joint action space $\mathbb{A}$, with components $a_i \in A_i, i : 1 \ldots n$. For every agent i, a distribution over possible rewards is associated with each joint action $\vec{a} \in \mathbb{A}$, i.e., the function $r_i : \mathbb{A} \rightarrow \mathbb{R}$ denotes agent i's expected payoff. The payoff function r_i is often called the *utility function* of the agent. It represents the preference relation each agent has on the set of action profiles. If $\sum_i r_i = 0$ for all joint actions $\vec{a}$, the game at hand is called a zero-sum game; otherwise, it is a non-zero-sum game. The tuple $(n, \mathbb{A}, r_{1\ldots n})$ thus defines a single-stage strategic game, also called a normal-form game. Usually, such a game is played repeatedly. In this case, agents can adopt either a pure strategy or a mixed strategy. In a pure strategy, a single action is always played, i.e., all probability lies on one action. In a mixed strategy, a set of actions is played with a certain probability per action.

By convention, two-agent strategic games are represented in a matrix form. An example is given in Figure 1. The rows and columns represent the actions for respectively agent 1 (the row player) and agent 2 (the column player). In the matrix the (expected) payoffs for every joint action can be found; the first (second) number in every table cell is the (expected) reward for the row (column) agent.

Nash equilibrium. The agents are said to be in a Nash equilibrium when it is the case that no agent can increase its own reward by changing its strategy when all the other agents stick to their strategy. In other words, there is no incentive for the agents to play any other strategy than their Nash equilibrium strategy. John Nash proved that every strategic game with finitely many actions has at least one pure or mixed Nash equilibrium (Nash, 1950a).

Pareto optimality. Despite this proof, describing an optimal solution for strategic games is not always easy. For instance, an equilibrium point is not necessarily unique. If more than one equilibrium point exists, they do not always give the same utility to the agents. Pareto optimality is a concept introduced to distinguish between multiple equilibria. The outcome of a game is said to be Pareto-optimal if there exists no other outcome for which all agents simultaneously perform better. Similarly, a joint action $\vec{a}$ is said to Pareto-dominate another joint action $\vec{a}'$ if all agents receive at least the same payoff with $\vec{a}$ as with $\vec{a}'$, and one agent receives strictly more.

The classical example of a situation where the individually rational action leads to inferior results for each agent, is the Prisoner's Dilemma Game (see Figure 2). This game has just one pure Nash equilibrium, i.e., joint action (a_{12}, a_{22}). However, it is not a Pareto-optimal equilibrium. Joint action (a_{11}, a_{21}) is obviously superior; it is the only strategy which Pareto-dominates the Nash equilibrium, but it is not a Nash equilibrium itself. Other games in which the individually rational action is not a 'good' one include the Traveler's Dilemma (Basu, 2007) and the Ultimatum Game, which will be discussed in detail later.

Bounded rationality. After introducing the concept of Nash equilibria, John Nash participated in a study with human subjects to assess their behavior in various games (Kalisch *et al*., 1952).

	a_{21}	a_{22}
a_{11}	$(1, -1)$	$(-1, 1)$
a_{12}	$(-1, 1)$	$(1, -1)$

Figure 1 The Matching Pennies Game. In this game, two agents independently choose which side of a coin to show to the other. If they both show the same side, the first agent wins, otherwise the second agent wins.

	a_{21}	a_{22}
a_{11}	$(5, 5)$	$(0, 10)$
a_{12}	$(10, 0)$	$(1, 1)$

Figure 2 The Prisoner's Dilemma Game. Two prisoners committed a crime together. They can either confess their crime (i.e., play the first action) or deny it (i.e., play the second action). When only one prisoner confesses, he takes all the blame for the crime and receives no payoff, while the other one gets the maximum payoff of 10. When they both confess, a payoff of 5 is received by both prisoners. Otherwise, they only get 1.

The study was considered a failure, because people failed to find the Nash equilibria. However, it turned out to be one of the first studies showing that (additional) information and fairness matter to humans. A few years later, Herbert Simon coined the term bounded rationality (Simon, 1957, 1972; Simon and Newell, 1972) to describe the phenomenon that most people are only partly rational, and are in fact emotional or even irrational in the remaining part of their actions. Bounded rationality suggests that people use heuristics to make decisions, rather than a strict rigid rule of optimization. They do this because of the complexity of the situations at hand, and the inability to process and compute all alternatives. Simon describes a number of dimensions along which 'classical' models of rationality can be made somewhat more realistic, while sticking within the vein of fairly rigorous formalization.

Evolutionarily stable strategies. Evolutionary Game Theory introduces another equilibrium concept than the Nash Equilibrium, namely the Evolutionary Stable Strategy (ESS) (Maynard-Smith, 1982; Tuyls and Nowe, 2005). This concept accounts for the stability of strategies in a population of agents. Imagine that such a population is playing the same strategy, i.e., every agent does the same. Assume that this population is invaded by a different strategy, which is initially played by a small number of the total population. If the reproductive success of the new strategy is smaller than the original one, it will not overrule the original strategy and will eventually disappear. In this case we say that the strategy is evolutionary stable against this new appearing strategy. More generally, we say a strategy is an ESS if it is robust against evolutionary pressure from any appearing mutant strategy. For a formal definition of an ESS and its relation to Nash Equilibria, we refer to Tuyls and Nowe (2005).

3 Descriptive models of human fairness

In this section, we discuss descriptive models of human fairness, i.e., models that describe why and how humans act in a fair way instead of choosing the individually rational solution. To this end, we describe three descriptive models of fairness and their explanantions for fair behavior: inequity aversion, priority awareness and reciprocal fairness. The inequity-averse and priority-aware models primarily aim at explaining the human concept of fairness, related to observed differences in payoff. Reciprocal models provide an explanation of how fairness may arise in situations where humans interact repeatedly, directly as well as indirectly.

3.1 Inequity aversion

In Fehr and Schmidt (1999), inequity aversion is defined as follows: '*Inequity aversion means that people resist inequitable outcomes; i.e., they are willing to give up some material payoff to move in the direction of more equitable outcomes*'. One of the first to discuss and model the role of inequity aversion in fairness is Matthew Rabin (1993). However, Rabin's model requires an explicit representation of fair intentions and is only applicable to two-agent, zero-sum games. The inequity-averse model developed by Fehr and Schmidt (1999) does not suffer from these drawbacks. Essentially, it assumes that a fraction of people are motivated by fairness considerations, in addition to being rational. It is known that this motivation is present in many people, even without previous reinforcement.

3.1.1 Homo Egualis

To model inequity aversion, an extension of the classical game theoretic actor is introduced, named Homo Egualis (Fehr and Schmidt, 1999; Gintis, 2001). Homo Egualis agents are driven by the following utility function:

$$u_i = x_i - \frac{\alpha_i}{n-1}\sum_{x_j > x_i}\left(x_j - x_i\right) - \frac{\beta_i}{n-1}\sum_{x_i > x_j}\left(x_i - x_j\right) \tag{1}$$

Here, u_i is the utility of agent $i \in \{1, 2, \ldots, n\}$. This utility is calculated based on agent i's own payoff, x_i, and two terms related to considerations on how this payoff compares to the payoffs x_j of other agents j: every agent i experiences a negative influence on its utility for other agents j that have a higher payoff as well as other agents that have a lower payoff. More precisely, all payoffs $x_j > x_i$ are first summed (using $\sum_{x_j > x_i}\left(x_j - x_i\right)$), subsequently weighed by a parameter α_i, and finally normalized with respect to the number of agents n. All payoffs $x_j < x_i$ are summed and normalized in a similar manner, but weighed by a parameter β_i. The two resulting terms are subtracted from the utility of agent i. Thus, given its own payoff x_i, agent i obtains a maximum utility u_i if $\forall j : x_j = x_i$. Research with human subjects provides strong evidence that humans care more about inequity when doing worse than when doing better in society (Fehr and Schmidt, 1999). Thus, in general, $\alpha_i > \beta_i$ is chosen. Moreover, the β_i-parameter must be in the interval [0, 1]: for $\beta_i < 0$, agents would be striving for inequity, and for $\beta_i > 1$, they would be willing to 'burn' some of their payoff in order to reduce inequity, since simply reducing their payoff (without giving it to someone else) already increases their utility value.

3.1.2 Inequity aversion in common games

The Homo Egualis utility function has been shown to adequately describe human behavior in various games, including the Ultimatum Game and the Public Goods Game. In this section, we will briefly discuss this. For a comprehensive overview, we refer to Fehr and Schmidt (1999) and Dannenberg *et al.* (2007). We additionally apply the Homo Egualis function to a new multi-agent extension of the Ultimatum Game and to the Nash Bargaining Game.

Ultimatum Game. The Ultimatum Game (Gueth *et al.*, 1982) is a simple bargaining game, played by two agents. The first agent proposes how to divide a (rather small, e.g., \$10) reward R with the second agent. If the second agent accepts this division, the first gets his demanded payoff and the second gets the rest. If however the second agent rejects, neither gets anything. The game is played only once, and it is assumed that the agents have not previously communicated, i.e., they did not have the opportunity to negotiate with or learn from each other. The individually rational solution (i.e., the Nash equilibrium) to the Ultimatum Game is for the first agent to leave the smallest positive payoff to the other agent. After all, the other agent can then choose between receiving this payoff by agreeing, or receiving nothing by rejecting. Clearly, a small positive payoff is rationally preferable over no payoff at all.

However, research with human subjects indicates that humans usually do not choose the individually rational solution. Hardly any first agent proposes offers that lead to large differences in payoff between the agents, and hardly any second agent accepts such proposals. Oosterbeek *et al.* (2004) analyze many available experiments with humans. It is indicated that the average proposal in the two-agent Ultimatum Game is 40%, with 16% of the proposals being rejected by the other agent. Our own experiments confirm this (De Jong *et al.*, 2008b). Cross-cultural studies of small cultures have shown that these numbers are not universal. However, independent of culture, the individually rational solution is hardly ever observed (Henrich *et al.*, 2004).

Using the Homo Egualis utility function with two agents, Fehr and Schmidt (1999) calculate that the optimal payoff for agent 1 depends on two factors, viz. β_1 and α_2. More precisely, in the two-agent game, we have $n = 2$ and $x_2 = R - x_1$. Thus, the Homo Egualis function can be rewritten for both agents as:

$$u_i = x_i - \alpha_i \max(R - 2x_i, 0) - \beta_i \max(2x_i - R, 0) \tag{2}$$

Thus, if $\beta_1 > 0.5$, agent 1's utility u_1 will decrease with values of $x_1 > 0.5R$, which implies that agent 1 will give $0.5R$ to agent 2 if $\beta_1 > 0.5$. If $\beta_1 < 0.5$, agent 1's utility is not decreased by increasing his payoff x_1. The agent would like to keep everything to himself. However, he must ensure agent 2 receives a payoff that is not rejected. Agent 2 will reject iff $x_2 - \alpha_2 (R - 2x_2) \leq 0$. Solving this inequality with respect to x_2, we obtain:

$$x_2 \geq \frac{\alpha_2}{1 + 2\alpha_2} \cdot R \tag{3}$$

Note that $\lim_{\alpha_2 \to \infty} = 0.5R$. Thus, agent 2 can expect to obtain at most half of the total reward. For additional clarity, the functional mapping between x_1 and u_1 is illustrated in Figure 3. From this figure, it is clear that the utility function for agent 1 is not continuous: there is a discontinuity immediately after the maximum.

Multi-agent Ultimatum Game. Usually, the Ultimatum Game is played with only two agents. As we are interested in a multi-agent perspective, we also analyze the role of inequity aversion in Ultimatum Games with more than two agents. There are various extensions of the Ultimatum Game to more agents, e.g., introducing proposer competition or responder competition. We propose a different extension. More precisely, we define a game in which $n - 1$ agents one by one take a portion of the reward R. The last agent, n, receives what is left. In this game, once again any payoff distribution in which $\forall\, i : u_i \geq 0$ holds, is acceptable. Since in general $\forall\, i, j : \alpha_i > \beta_j$, we can derive that the agent that receives the lowest payoff, also has the lowest utility. More precisely, this agent's utility value can be calculated as $u_i = x_i - \frac{\alpha_i}{n-1} \sum_{x_j > x_i} (x_j - x_i)$, since the term involving β_i evaluates to 0 for the agent with the lowest payoff. Thus, assuming a reward of R, we calculate:

$$\begin{aligned} u_i = x_i - \frac{\alpha_i}{n-1} \sum_{x_j > x_i} (x_j - x_i) &\geq 0 \\ x_i - \frac{\alpha_i}{n-1} \left[\sum_{x_j > x_i} x_j - \sum_{x_j > x_i} x_i \right] &\geq 0 \\ x_i - \frac{\alpha_i}{n-1} [R - x_i - (n-1)x_i] &\geq 0 \end{aligned}$$

This simplifies to:

$$x_i \geq \frac{\alpha_i}{\alpha_i n + n - 1} \cdot R \tag{4}$$

For instance, with three agents and $\alpha_3 = 0.6$, we obtain that the worst-performing agent needs to obtain at least $0.1578R$ in order to accept the deal at hand. As long as the other agents obtain more, they will accept any deal. Typically, as in the two-player Ultimatum Game, the last agent will be the worst-performing agent, as the other agents can actively reduce his payoff to the lowest value that still leads to a positive utility.

Nash Bargaining Game. The Nash Bargaining Game (Nash, 1950b) is traditionally played by two agents, but can easily be extended to more agents. In this game, all agents simultaneously determine how much payoff x_i they are going to claim from a common reward R. If $\sum_i x_i > R$, everyone receives 0. Otherwise, everyone receives what they have asked for. Note that payoffs may not sum up to R, i.e., a Pareto-optimal solution is not guaranteed. The game has many Nash equilibria, including one where all agents request the whole R.

The common human solution to this game is an even split (Nydegger and Owen, 1974; Roth and Malouf, 1979; Yaari and Bar-Hillel, 1981; Van Huyck *et al.*, 1995). Inequity aversion may increase the ability of agents to find such a fair solution. To this end, we give agents an additional action, that is, even if the payoff distribution was successful, agents may compare their payoff with that of others. Then, if their payoff is too small, they may reject the distribution, once again leading to all agents obtaining a payoff of 0. To decide whether their payoff is satisfactory, agents

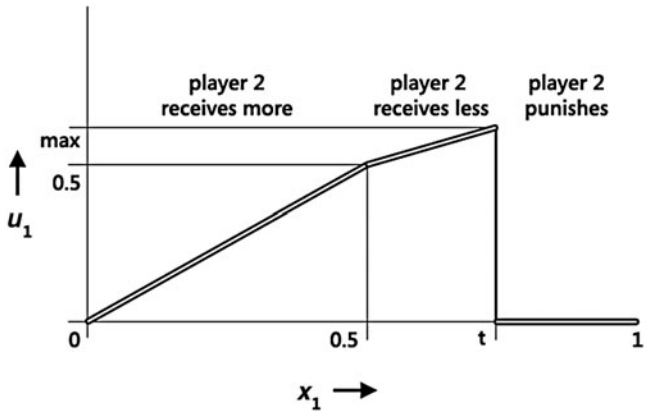

Figure 3 Homo Egualis in the two-agent Ultimatum Game. We illustrate the functional mapping between the payoff agent 1 keeps to himself (x_1) and the utility experienced by this agent (u_1). Agent 2 can reject in case of a negative utility, i.e., if the payoff agent 1 keeps to himself exceeds a threshold $t = \frac{\alpha_2}{1+2\alpha_2} R$ In this case, both agents receive 0.

use the Homo Egualis utility function, as in the Ultimatum Game. Thus, we can perform the same analysis as in the Ultimatum Game and obtain that any solution for which every agent obtains at least $\frac{\alpha_i}{\alpha_i n+n-1} \cdot R$ is not rejected. For example, with $n = 2$ and $\alpha_1 = \alpha_2 = 0.6$, every agent should obtain at least $0.27R$.

Public Goods Game. In the Public Goods Game (Fehr and Schmidt, 1999; Dannenberg *et al*., 2007), every agent receives an amount of money, (part of) which can be invested in a common pool. Every agent simultaneously chooses how much to invest. Then, everyone receives the money he kept to himself, plus the money in the common pool, multiplied by a certain factor (usually 3) and divided equally among the agents. To gain the most profit, everyone should cooperate and contribute their whole private amount. However, every agent can gain from solely not contributing. Thus, the Nash equilibrium is for every agent to defect, leading to a monetary payoff that is (usually) 3 times lower than the optimal payoff. Sigmund[2] compares this game to hunting a mammoth in prehistoric times: every hunter runs the risk of being wounded by the mammoth, but if they all work together, they can easily catch the animal. The game is one of the many examples of the tragedy of the commons (Aristotle, 1885; Hardin, 1968), in which Adam Smith's famous 'invisible hand' (Vaughn, 1987) fails to work.[3] Human behavior in this game is rather surprising: initially, many people cooperate, but in repeated games, this cooperation decreases. The initial success obtained by few defectors is quickly learned by the other agents, leading to lower and lower payoffs. Introducing the option to punish defectors leads to cooperation, but as we will discuss in Section 3.3, punishing is not a dominant strategy.

Dannenberg *et al*. (2007) performed an extensive study to assess whether inequity aversion can be used to explain this human behavior. The study is based on an idea coined by Fehr and Schmidt (1999), i.e., in the paper that introduced inequity aversion. Subjects were first presented with modified Ultimatum Games to group them into classes with similar preferences (i.e., similar estimated α_i- and/or β_i-parameters). Next, subjects were matched into pairs in three different ways, i.e., 'fair' pairs with highly inequity-averse subjects, 'selfish' pairs with subjects who did not care strongly about inequity, and 'mixed' pairs in which one subject was 'fair' and the other one 'selfish'. Next, pairs played a standard Public Goods Game and one in which punishment could be used. Results indicate that the inequity-averse model at least has some explanatory

[2] In a keynote talk at the European Conference on Complex Systems 2007.

[3] Smith claimed that, in a free market, an individual pursuing his own self-interest tends to also promote the good of his community as a whole through a principle that he called 'the invisible hand'. He argued that each individual maximising revenue for himself maximises the total revenue of society as a whole, as this is identical with the sum total of individual revenues.

power, especially in the 'fair' pairs. The parameter that seems most relevant is the one modeling advantageous inequity (i.e., β_i). The fact that inequity aversion can be used to explain behavior in Public Goods Games is rather surprising, since the model does not take into account considerations such as reciprocity, intentions or even just repeated games.

In conclusion, we see that the inequity-averse Homo Egualis model 'is able to explain an impressive amount of experimental evidence not in line with the standard model of selfish behavior' (Dannenberg *et al.*, 2007). However, it should be noted that there are also experiments in which human behavior is not adequately captured by a utility model that is exclusively based on inequity aversion and material interest (Charness and Rabin, 2002). In many cases, subjects are more concerned with increasing social welfare than with reducing differences in payoffs. Subjects are also motivated by additional information they may have about each other, and by reciprocity: they become less cooperative in the presence of defectors and sometimes punish unfair behavior. This leads to two other models, viz. priority awareness and reciprocal fairness, which will be outlined in the following two sections.

3.2 Priority awareness

In our own work (De Jong *et al.*, 2006, 2008b), we observed that people's concept of a fair deal depends on additional information they may have about themselves and other people. For example, everybody can agree that priority mail should be delivered faster than regular mail, since priority mail is more expensive. Similar issues arise in many common applications of multi-agent systems and should therefore be addressed when looking at fairness. Inequity aversion might not be sufficient to describe fair deals for which additional information should be considered. In real-world examples, the nature of the additional information can vary. Examples include wealth, probabilities of people belonging to a certain group or priorities involved in the task at hand. In our work, we assume that the additional information can be expressed with one value per agent, i.e., the *priority*. Currently, we assume that the priority values are true and are known by all agents.

3.2.1 Priority awareness in humans

We performed a series of experiments to study the human concept of fairness in priority problems. We present an overview here.

An initial experiment: the vegetable shop. In previous research, we were developing a decentralized multi-agent system for resource gathering in storage facilities (De Jong *et al.*, 2006). Since the system was intended to work on behalf of customers, we wanted to incorporate a 'human' performance measure, as opposed to other work (e.g., Weyns *et al.*, 2005), which used analytical measures such as a low average waiting time. To this end, we developed a survey (see Figure 4), in which 50 faculty members and students participated. Respondents were asked to place a robot between two pickup points in such a way that 'all customers will be satisfied', given that there was a 0.6 probability that a customer would request an item stored at (A), and a 0.4 probability that an item stored at (B) would be requested. Surprisingly, 45 of the 50 respondents did not

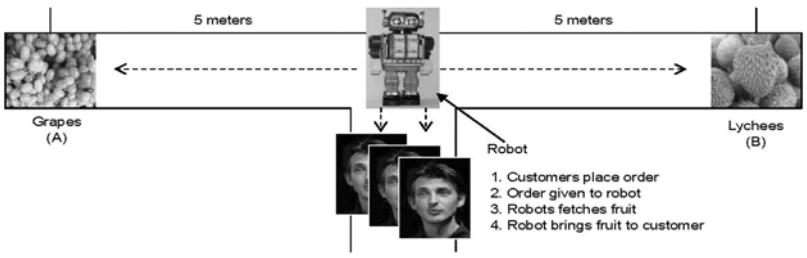

Figure 4 Priority awareness in a vegetable shop. Respondents are asked to place the robot on the line between (A) and (B), in such a way that all customers of the shop are satisfied. The probability that customers order the item located at (A) is varied. As a result, respondents select different positions for the robot.

choose a rational solution, such as minimizing the average waiting time (by placing the robot at (A), see Section 3.2.4). Instead, the robot was placed somewhere between (A) and the middle. When we increased the probability that an item stored at (A) would be requested, the robot was placed more to the left (i.e., closer to (A)), but only for very high probabilities was (A) chosen as the location for the robot. Thus, it seemed that our respondents favored the requesters of the more popular item with a slight, but not too big, advantage over the requesters of the less popular item.

A prioritized Ultimatum Game. After some initial experiments (such as the one described above), we created a larger, more structured experiment, after the example of many experiments with human fairness in two-player Ultimatum Games (for an overview of such experiments, see Fehr and Schmidt, 1999; Oosterbeek *et al*., 2004). We asked students and staff members of three faculties of the University of Maastricht to participate in our Ultimatum-Game survey, which was developed in cooperation with an experimental psychologist.[4] Respondents were asked various control questions and a number of Ultimatum Game dilemmas, in which they had to indicate how much they would offer (as a first player) or accept at least (as a second player). In the end, 170 surveys were submitted, of which 160 were usable. Of the 160 respondents, 38 were familiar with the Ultimatum Game; the remaining 122 were not.

To introduce the notion of priorities more explicitly in the Ultimatum Game, we varied two quantities. First, after playing some standard Ultimatum Games, participants were told that the other player was 10 times poorer or 10 times wealthier than they were. In this case, people could either be fair to poorer people (i.e., give them more) or exploit poorer people (i.e., give them less, because they will accept anyway). If priorities mattered to the respondents, they would go for the fair option, since clearly a poorer player needs the money more. Second, the amount of money that had to be divided was varied between EUR. 10, EUR. 1 000 and EUR. 100 000, to determine whether this had any effect on people's attitude with respect to poorer, equal or richer people. We will now present the most important results, which are graphically represented in Figure 5. Note that the answers concerning the second player have not been included: in all cases, respondents answered consistently for the first and second player, meaning that they always offered at least the amount they personally would accept.

Since we were asking people to imagine they had to divide money, instead of giving them actual money to divide, we first needed to assess whether this difference had an important impact on the results. This, fortunately, was not the case, as the behavior we found was in line with behavior found earlier (Fehr and Schmidt, 1999; Oosterbeek *et al*., 2004). For instance, many people were willing to give away 50%, and some people offered less. With an increasing amount, we see that the first player keeps more to himself and that this behavior is accepted by the second player. This is rather obvious; after all, it is a lot more difficult to say no to 10% of EUR. 100 000 than to 10% of EUR. 10. Interestingly, most literature on this subject states that the amount at stake does *not* influence the offered (and accepted) proportion (e.g., see Fehr and Schmidt (1999); a notable exception is Zollman (to appear)). However, experiments with high amounts at stake were mostly performed in relatively poor countries, simply because research institutes cannot afford to let people play games worth EUR. 100 000 in Europe. Perhaps, this explains why such 'high-stake' games have such a fair outcome in general.

Players' behavior in the normal Ultimatum Game can already be successfully explained using the Homo Egualis utility function, as outlined above. However, the results of our survey show that priorities, which are not explicitly modelled using Homo Egualis, strongly matter to human players of the Ultimatum Game. Before confronting participants with Ultimatum Games in which agents had unequal wealth, we asked them whether they had thus far assumed that the other player was poorer, wealthier or equally wealthy. Ninety-two percent of the participants assumed that the other player was equally wealthy; the remaining 8% was almost equally divided between the other two options. Subsequently, the participants were confronted with games in which the

[4] The survey can still be viewed at `http://www.cs.unimaas.nl/steven.dejong/survey`.

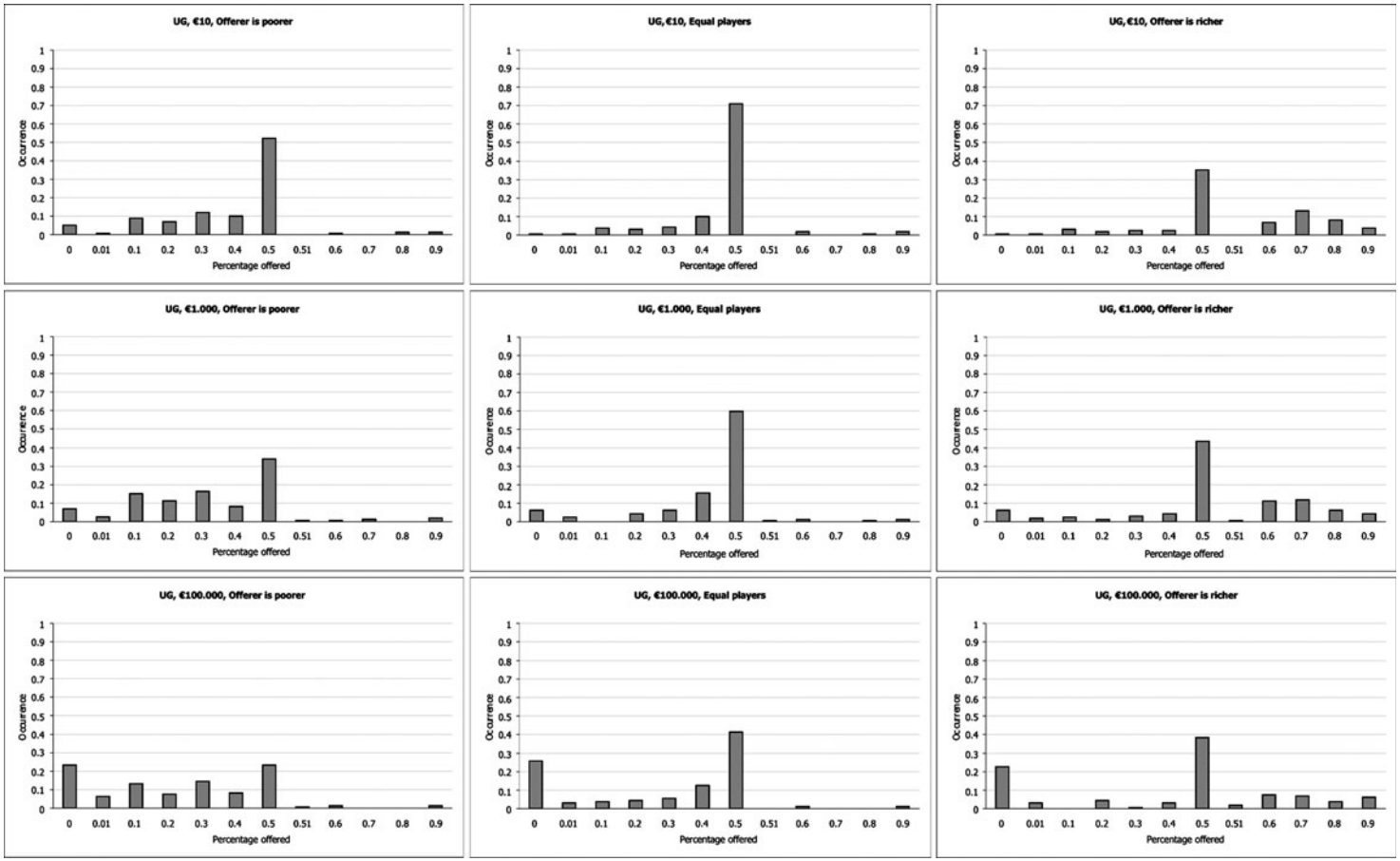

Figure 5 Offers made by the first agent in the Ultimatum Game, depending on the amount at stake and the relative wealth of the other agent. Offers are binned in 12 bins, of which the *lower bound* is displayed in every chart

other player was 10 times more or less wealthy than they were. Results indicate that people are actually fair to poorer people, and expect the same in return from richer people. Moreover, poorer opponents were given substantially more money than equal opponents, richer opponents were given substantially less (as can be seen in Figure 5). This indicates that priorities matter to humans in this game. Once again, all participants were willing to accept their own offers.

3.2.2 The need for a new model

Given the three experiments above, we see that priorities indeed matter to human agents. This is not sufficiently described by the Homo Egualis function. Most notably, there are two problems. First, as clearly indicated in Fehr and Schmidt (1999) (p. 850), the Homo Egualis actor does not model the phenomenon that people actually like to be better than other people. In priority problems, people actually accept inequity to a certain extent: agents with a high priority should be allowed to obtain a higher payoff than agents with a low priority. Priorities would have to be encoded in an indirect way using the model's parameters (most notably α_i), which then needed to be changed with respect to every possible other agent. Second, the significant number of participants who offer more than half of the amount at stake to a poor second agent in prioritized Ultimatum Games, is not described by Homo Egualis at all. The priority-aware model has to be able to capture these phenomena in a straightforward way, using the notion of priorities.

3.2.3 The priority awareness model

We model people's perception of fairness in the presence of additional information using a descriptive fairness model called *priority awareness*. In real-world examples, the nature of the additional information can vary. Examples include wealth, probabilities of people belonging to a certain group or priorities involved in the task at hand. In our model, we assume that the additional information can be expressed with one value per agent i, i.e., the *priority* p_i. Moreover, every agent attaches an individual *weight* $w_i \in [0, 1]$ to its priority. With $w_i = 0$, agent i does not pay attention to its priority at all and with $w_i = 1$, the priority matters strongly. As we also saw when we developed our conceptual model (De Jong *et al.*, 2008b), people usually are rather tolerant, i.e., their boundary between verdicts such as 'acceptable' and 'not acceptable' is not immediate. For instance, if someone wishes to receive 50% of the reward, he will probably also accept 49.5%, less probably accept 49% etc. To model this phenomenon, we introduce a weight interval: the value of w_i does not have to be specified exactly, but may be in a range; $w_i \in [w_{i_{\min}}, w_{i_{\max}}]$.

Since the inequity-averse Homo Egualis model already describes human behavior in a variety of games in which priorities are not present, we propose an extension to this model for games that do include priorities. More precisely, we introduce the notion of a *perceived payoff* x_i, which is defined as:

$$\hat{x}_i = \frac{w_i}{p_i} x_i + (1 - w_i) x_i \tag{5}$$

Thus, if agent i has a high priority (and $w_i \neq 0$), he has a lower perceived payoff $\hat{x}_i$ than the payoff x_i he is actually receiving. The bounds on the parameter w_i prevent agents from having an irrealistic perceived payoff. For instance, with $w_i < 0$, agents with a low priority would perceive a low payoff as insufficient. The perceived payoffs can be used in the modified Homo Egualis utility function to determine whether a certain payoff distribution $(x_1, \ldots, x_n)$ is acceptable, given the agents' priority and weight parameters, i.e.:

$$u_i = \hat{x}_i - \frac{\alpha_i}{n-1} \sum_{\hat{x}_j > \hat{x}_i} \left(\hat{x}_j - \hat{x}_i\right) - \frac{\beta_i}{n-1} \sum_{\hat{x}_i > \hat{x}_j} \left(\hat{x}_i - \hat{x}_j\right) \tag{6}$$

Note that this means that agents may need to know or estimate each others' perceived payoffs $\hat{x}_i$, in addition to the parameters α_i and β_i. However, with various cryptographic techniques it is possible to allow agents to calculate their utilities without explicitly knowing the values of priorities and payoffs of the other agents (Denning, 1982).

The concept of equity (or inequity) is now viewed from the perspective of a transformed perception of the payoffs. Thus, for instance, two agents with $p_1 = 1$, $I_2 = 2$ and $w_1 = w_2 = 1$ perceive an

equal split in case agent 2 receives an actual payoff that is twice that of agent 1. In an interaction between n agents, a certain payoff distribution is acceptable if every agent i can assign a weight w_i from the range $[w_{i_{\min}}, w_{i_{\max}}]$ to his priority p_i, in such a way that he experiences a positive utility. Thus:

$$\forall i \exists w_i \in [w_{i_{\min}}, w_{i_{\max}}] : u_i > 0 \rightarrow accept \tag{7}$$

Otherwise, one or more agents have reason to reject. Note that in this case, the agent with the lowest perceived payoff is the first to reject. This is not necessarily the same agent as the agent with the lowest actual payoff.

3.2.4 Describing experimental results

Unlike the inequity-averse model, the priority-aware model can be used to explain all human behavior in the various experiments we performed. We will discuss them here.

The vegetable shop. In the vegetable shop experiment (see Figure 4), we determined the human response to a very simple priority problem, in which customers of a (rather impractical) shop can be divided in two groups. Each group would like to buy a different item, located at (A) and (B) in the shop, respectively. A robot has to fetch items, but has to be placed somewhere between (A) and (B) such that the delays experienced by the customer groups are balanced. We model each of the groups as a single agent (respectively A and B), and assign to the (two) agents a priority value p_i equal to the probability that a customer belongs to their associated group. Let us assume for easier calculation that the distance between locations (A) and (B) is 2 and that the customers wait precisely in the middle between these locations. Thus, if the robot is positioned at $a \in [0, 2]$, customers represented by agent A will experience a delay $d_A = a + 1$ before the robot delivers the requested item. Customers represented by agent B will experience a delay of $d_B = 2 - a + 1 = 3 - a$. Since waiting time can be seen as negative payoff, we define the payoffs of the agents as $x_A = 3 - a +$ and $x_B = a + 1$. Thus, A obtains a higher payoff with lower values of a, and the other way around. The goal of the experiment is now to find a position a that satisfies both agents A and B, i.e., a fair position. More precisely, since both agents should be treated equally well, we would like both agents to have the same utility, i.e., $u_A = u_B$.

In case the group of customers wishing to obtain the item at (A) is equal in size to the group of customers wishing to obtain the item at (B), we obtain $p_A = p_B = 0.5$. With equal priorities for all agents, the priority-aware model is identical to the inequity-averse one, i.e., we can set $\hat{x}_A = x_A, \hat{x}_B = x_B$. Let us assume that we choose an $a \leq 1$ (for $a > 1$, everything can simply be inverted). In this case, we obtain $\hat{x}_A \geq \hat{x}_B$. Now we can calculate the utility values as $u_A = \hat{x}_A - \beta_A(\hat{x}_A - \hat{x}_B) = 3 - a - \beta_A(2 - 2a)$ and $u_B = \hat{x}_B - \alpha_B(\hat{x}_A - \hat{x}_B) = 1 + a - \alpha_B(2 - 2a)$. As has been mentioned above, the fair position satisfies $u_A = u_B$. This is only possible if the robot is placed in the middle, i.e., $a = 1$ (or if, coincidentally and ignorably, $\beta_A + \alpha_B = 1$). This result corresponds exactly to human behavior.

For a situation in which the priorities differ, priority awareness is no longer identical to inequity aversion. The latter model would predict the same outcome as with equal priorities, i.e., a=1. With priority awareness, however, we can use the same analysis as above: we simply have to find the fair position for which $u_A = u_B$. For instance, assume that there is a 60% chance for customers to request the item at (A). Thus, we set the priorities to $p_A = 0.6$ and $p_B = 0.4$. Now, for $w_A = w_B = 1$, the perceived payoffs are $\hat{x}_A = \frac{3-a}{0.6}$ and $\hat{x}_B = \frac{1+a}{0.4}$. For $u_A = u_B$ to hold, we must have $\hat{x}_A = \hat{x}_B$ (or once again a coincidental, ignorable relation between the various parameters), and for this, we must have $a = 0.6$. For lower priority weights w_A and w_B (which are possible in our model), a can become larger, up to $a = 1$ for $w_A = w_B = 0$. This outcome corresponds nicely to what our human subjects did: we saw that most human subjects placed the robot at $a \in [0.6, 0.8]$, with 0.6 being the most frequent choice.

The prioritized Ultimatum Game. The priority-aware model can also be used to explain or predict human strategies in prioritized Ultimatum Games. As has been mentioned before, in the absence of priorities, the model is equivalent to the inequity-averse model. Thus, human strategies emerging in regular Ultimatum Games are explained in the same (successful) way by both models.

In Figure 5, we see the median offer is 50% here, with the average varying from 45% (sharing EUR. 10) to around 25% (sharing EUR. 100 000). Offers of more than 50% are very rare. This is indeed predicted by the inequity-averse model (as well as the priority-aware model), as we have seen earlier. Using data gathered in these Ultimatum Games, we can estimate the participants' α_i- and β_i-parameters for every amount at stake, in a similar way as described by Dannenberg *et al.* (2007). Clearly, they decrease with an increasing amount. In conclusion, people become more and more selfish with an increasing amount, and are not punished for that.

Introducing priorities, i.e., richer or poorer players, has no effect on the predictions of the inequity-averse model, assuming that the α_i- and β_i-parameters of individual participants do not change. In Figure 5 however, we do see different strategies emerge. With the first (i.e., offering) player being poorer, the median offer is still 50%, but the average drops: around 30% (EUR. 10) to 15% (EUR. 100 000). Once again, players accept their own offers. This outcome can be predicted by the priority-aware model: assuming that $p_1 > p_2$, the model predicts that the first player wishes (and is allowed) to keep more to himself, which indeed happens. With the first player being richer, the difference between the inequity-averse and priority-aware model becomes even more clear. Many players are now willing to give more than half of the amount to the second player. The median offer is still 50%, but the average increases to around 65% (EUR. 10) and 45% (EUR. 100 000). This cannot be explained by inequity aversion alone, but priority awareness helps: using this model, we obtain that the *perceived* amount that player 1 gives to player 2 cannot be more than 50%, but the actual amount can. For instance, assume that player 1 has a low priority compared to player 2 (say, $p_1 = 1$ $p_2 = 4$). In this case, with the priority weights set to 1, a perceived 50–50% payoff distribution corresponds to an actual payoff distribution of 20–80%. Thus, it is very possible to give away more than half the actual amount.

In conclusion, we see that priority awareness helps explaining human strategies in case a reward has to be divided among people for which this reward is of different importance.

3.3 Reciprocal fairness and altruistic punishment

The most important limitation of the inequity-averse and priority-aware models is that they do not explicitly explain how fair behavior evolves with repeated interactions between agents (Fehr and Schmidt, 1999).[5] More precisely, a group of people repeatedly playing the same game may start by playing in an individually rational manner, but for some reason may end up playing in a fair, cooperative manner (or the other way around). Reciprocal fairness models aim at providing an answer to the questions why and how this happens. The main idea is that humans cooperate because of direct and indirect reciprocity—here, direct means that a person is nice to someone else because he expects something in return from this other person, and indirect means that an agent is nice to someone else because he expects to obtain something from a third person. It turns out that the opposite, i.e., being nasty to someone who was nasty to you (i.e., punishment), has an even greater effect on cooperation (Sigmund *et al.*, 2001). However, being nasty may be costly, and thus, one would expect that humans only punish when they are sure to encounter the object of punishment again. This is not the case: even in one-shot interactions, humans consistently apply punishment if this is allowed. Since this is not of direct benefit to the punisher, this phenomenon is referred to as altruistic punishment (see, e.g., Yamagishi, 1986; Fehr and Gaechter, 2000, 2002). Many researchers argue that altruistic punishment only pays off when the reputation of the players somehow becomes known to everyone (Milinski *et al.*, 2002, Fehr, 2004). There are also alternative explanations such as volunteering (Hauert *et al.*, 2002, 2007) or fair intentions (Falk and Fischbacher, 2006). Moreover, researchers have found physical (i.e., neuronal or genetical) explanations for the fact that humans (and other social species, such as ants) are altruistic. We will provide an overview here.

[5] Using priority awareness, this may be possible if we allow agents' priorities to change over time.

Physical explanations. In humans, we see that there are many moral and justicial pressure devices. Humans take the law in their own hands in the absence of those devices. Recent studies show a physical basis for this behavior, i.e., our tendency to be socially aware seems to be explicitly encoded in our neurons. For instance, Sanfey *et al*. (2003) studied the brain activity of humans playing Ultimatum Games. They perceived strong activity in brain areas related to both cognition and emotion. Moreover, Knoch *et al*. (2006) show that the right prefrontal cortex plays an important role in fair behavior. Disrupting this brain area using (harmless) magnetic stimulation made subjects significantly more selfish.

Apart from a neuronal explanation, researchers also found a genetical explanation for reciprocal fairness. Although genes are commonly regarded as being selfish, i.e., purely focused on successful replication (Dawkins, 1976), a gene that promotes altruism might experience an increase in frequency if the altruism is primarily directed at other (potentially unrelated) individuals who share the same gene (Hamilton, 1964). Obviously, for this kin selection to actually work, the presence of the gene must be somehow perceivable and recognized. Hypothesizing that altruistic individuals might for instance have a green beard, Dawkins coined the term *green-beard effect*. Interestingly, in analogy to cooperators being vulnerable to exploitation by defectors, green-beard genes are vulnerable to mutant genes arising that produce the perceptible trait (i.e., the green beard) without the associated altruistic behavior. After being hypothetical for over 20 years, the first green-beard gene was indeed found in nature (Keller and Ross, 1998). In humans, similar (groups of) genes may physically control our urge to perform altruistic punishment. These genes apparently have had an evolutionary advantage.

Altruistic punishment. As has been mentioned above, researchers have studied how people are driven toward a fair solution even if this is not the optimal solution from an individually rational perspective. Punishment seems to be the driving force here. Altruistic punishment has been studied extensively using social dilemmas. There are two types of social dilemmas, viz. (1) social dilemmas in which a common resource needs to be distributed over a group of agents, as for instance the Ultimatum Game, and (2) social dilemmas in which every agent can invest a certain amount in order to obtain a larger returned payoff (given that everyone invests), as for instance in the Public Goods Game. In both cases, defection is the individually rational solution. In the first case, an individually rational agent will keep as much as possible to himself, knowing that the others can then choose between obtaining a very small payoff or nothing at all. In the second case, the individually rational solution is to defect by refusing to invest: if a single agent in the group follows this strategy, he can keep his original investment and receives a share in the investments of others. In contrast, cooperation is the optimal solution in both cases. In the first case, we see that low offers are considered 'rude' and will be rejected if this is possible, leading to zero payoff for everyone. In the second case, there is a large gain for everyone if everyone cooperates. The human response to the two types of social dilemmas is roughly the same: fair, optimal solutions emerge in the presence of punishment, and are not maintained in the absence of punishment. For instance, what we see in the Public Goods Game (without punishment) is that initially, many people cooperate, but in repeated games, this cooperation decreases. The initial success obtained by few defectors is learned by the others, leading to lower and lower payoffs.

Modeling this phenomenon, researchers used evolutionary game theory and population dynamics (Gintis, 2001; Tuyls and Nowe, 2005). Simple learning rules explain how people converge to defection. For instance, people may imitate others with a probability equivalent to their success. As a result of the rules of the Public Goods Game, we see that defectors obtain $\frac{rcN_c}{N_c+N_d}$ and cooperators obtain $\frac{rcN_c}{N_c+N_d} - c$, where r is the ratio the common good is multiplied with (typically 3), c is the contribution of cooperators (which is assumed to be constant), N_c is the number of cooperators and N_d the number of defectors. We see that defectors obtain a higher payoff, so they have a higher probability of being imitated. In the end, the whole population defects, even though in this game (and other examples of the tragedy of the commons), obtaining a fair, cooperative solution is in the material advantage of everyone.

The effect of introducing costly punishment in the Public Goods Game has been extensively studied (e.g., Yamagishi, 1986; Fehr and Gaechter, 2002). More precisely, participants can give

a sum of money to the experimenter to decrease the payoff of a defector. So, defectors lose $-\beta N_p$ (with N_p the number of punishers) in comparison to the non-punishing game and punishers lose $-\gamma N_d$ on their payoff (β and γ are constants). With this mechanism in place, cooperation increases, even if the participants are told that they will never interact with the same other participant(s) again. Thus, the punishment here is truly altruistic. Interestingly, Henrich *et al.* (2006) show that punishment and altruism are indeed correlated in human societies: more altruistic societies punish more, *et vice versa*.

Although this mechanism increases cooperation, it is not evolutionarily stable. Cooperators cannot invade a population of defectors, since a single cooperator would see its payoff reduced to (near-)zero. In addition, we can now have so-called second-order free-riders. Since punishment is costly, a participant may choose to cooperate, but at the same time, can refuse punishing defectors (i.e., in the first-order social dilemma, the participant chooses to cooperate, but in the second-order social dilemma, he actually defects by not wanting to punish). Thus, cooperators easily invade a group of punishers; after that, defectors can invade the cooperators. Higher-order punishment can be introduced and works from a computational point of view, but is not observed in humans. Thus, the question seems to shift: instead of studying why people learn to cooperate, we have to study why people actually perform altruistic punishment. As Boyd and Mathew (2007) note, this question has two parts from an evolutionary perspective. The first part is concerned with the question how contributors who punish can avoid being replaced by second-order free-riders. There has been much work on this topic lately, and plausible solutions have emerged (Henrich, and Boyd, 2001; Milinski *et al.*, 2002; Boyd *et al.*, 2003; Henrich, 2004; Panchanathan and Boyd, 2004), most notably related to reputation—humans tend to be cooperative towards someone they trust. However, these solutions still do not address the second part of the question, i.e., why punishment becomes established within populations in the first place.

Volunteering. A possible answer to this question is the concept of volunteering (Hauert *et al.*, 2002, 2007). If we once again see the Public Goods Game in analogy with hunting a mammoth, we see that the hunters may have volunteered to participate in the hunt. They could also have chosen to stay home safely, and collect mushrooms. Obviously, the people that collect mushrooms experience a lower food quality than the hunters, but they take this for granted, since picking mushrooms is also a lot less risky than trying to catch a mammoth. Also, since the hunters have volunteered to participate, they can expect that all hunters cooperate. To model this, an alternative third strategy can be introduced (with the first two being 'defect' and 'cooperate', but no punishment, as above), which could be labelled 'refuse'. Refusers obtain a payoff $s \in [0, (r-1)c]$. Thus, the payoff is higher than what agents obtain when everyone defects (i.e., 0), but at the same time, it is lower than what they obtain when some (or all) agents cooperate (i.e., at most rc). Thus, refusers can invade a population of defectors. Then, since s is lower than the payoff that would be obtained if everyone cooperates, rare cooperators can invade a population of refusers. Finally, defectors can once again invade the cooperators. These strategies oscillate endlessly (Hauert *et al.*, 2002). Boyd and Mathew (2007), show that adding punishment here, i.e., a fourth strategy 'punish' (which implies cooperation), makes it possible for punishers to invade the oscillating mixture of cooperators, defectors, and refusers, and once they do they tend to persist. As Boyd and Mathew (2007) note, this means that the population spends most of the time in a happy state in which cooperation and punishment of defectors predominate. Additionally, Hauert *et al.* (2002) and Nowak *et al.* (2004) find interesting results in finite populations. More precisely, if agents do not change their strategies, there will be four absorbing states, representing the four pure strategies. With small mutation rates, a stationary distribution is obtained. With a larger selection strength, punishing is the dominant strategy. This does not happen if the refusers are removed. In that case, defection is the dominant strategy. This is an interesting example of the problem of irrelevant alternatives.[6]

[6] This problem is described in many humorous anecdotes, such as one in which a man in a restaurant is trying to choose between the two dishes of the day (say, souvlaki and spaghetti), ultimately selects the spaghetti, but is then told by the waitress that there will also be curry today. The man responds by saying: 'This changes everything! I'd like souvlaki please.'

Intentions. In (Falk and Fischbacher, 2006), an alternative theory of reciprocity is presented, based on the concept of fair or kind intentions. In this theory, a reciprocal (i.e., fair) action is modeled as the behavioral response to an action that is perceived as either kind or unkind. The two central aspects of the model are the consequences of the action and the actor's underlying intentions. Several experimental studies suggest that fair intentions play a major role for the perception of kindness. For instance, human second players in the Ultimatum Game tend to punish more if the payoff is offered willingly by the first player, than if this player is forced to offer a randomly generated payoff (e.g., by performing a dice roll). Similarly, if the first player can only choose between the strategies 'give away 80%' or 'give away 20%', offers of 20% are rejected much less often. Inequity aversion is not able to explain this, but the kindness theory is. The theory is applied to various other games, and is shown to predict human behavior well (Falk and Fischbacher, 2006).

4 Computational models of human fairness

The descriptive models presented in the previous part of this paper may serve as a useful inspiration for prescriptive or computational models, i.e., models that drive the agent system in the direction of a cooperative or fair solution. In this section, we will discuss such contributions to prescriptive modeling of human fairness. We use the same order as in the previous section, i.e., we first discuss the development of an inequity-averse multi-agent system, which can learn valid solutions to the Ultimatum Game and the Nash Bargaining Game. The same system may be used to learn priority-aware solutions. Next, we address reciprocal multi-agent systems; many researchers have come up with reciprocal mechanisms to enhance cooperation and fairness in multi-agent games. Finally, we briefly discuss various other ideas closely related to computational modeling (but not necessarily to fairness), varying from mechanism design and computational social choice to normative systems.

4.1 Inequity-avert learning agents

In this section, we summarize our own work in the area of fairness in multi-agent systems (De Jong *et al*., 2008a). We use the inequity-averse model of fairness, which has been shown to adequately describe human behavior in the Ultimatum Game, the Nash Bargaining Game, and (to a degree) the Public Goods Game, as has been explained above. We apply this model in an adaptive multi-agent system, in which agents are learning by means of learning automata. The agents are then confronted with various Ultimatum and Nash Bargaining Games. It is important to note (and repeat) that purely rational agents obtain an unsatisfactory payoff at least in the Ultimatum Game, and that the Nash Bargaining Game has not yet been considered in the context of inequity aversion. We aim at answering two questions, viz. (1) do our agents learn to find and maintain valid solutions in both games, and (2) do these solutions correspond to solutions found by humans, as reported in literature?

4.1.1 Methodology

In our approach, we use a combination of the Homo Egualis utility function and Continuous Action Learning Automata (CALA). As has been described above, Egualis provides the necessary connection between the artificial agents and the human way of thinking in various settings. CALA facilitate the learning process. We will now discuss our architecture in more detail.

Learning automata. Originally, learning automata were developed for learning optimal policies in single-state problems with discrete action spaces (Narendra and Thathachar 1989; Tuyls and Nowe, 2005; Verbeeck *et al*., 2007). An automaton is assumed to be situated in an environment, in which it executes a certain action x from its discrete set of possible actions $\mathbb{A}$. This action x is observed by the environment and leads to a feedback $\beta(x)$ to the automaton. The automaton uses this feedback to update the probability that action x will be chosen again. Thus, a learning automaton is a simple reinforcement learner. With multiple (i.e., n) learning automata, every automaton i receives

feedback $\beta_i(\bar{x})$, resulting from the joint action $\bar{x} = (x_1, \ldots, x_n)$, but is not informed about the actions of the other automata. Nonetheless, with certain update schemes, learning automata have been shown to converge to an equilibrium point, e.g., a Nash equilibrium (Narendra and Thathachar, 1989).

Continuous Action Learning Automata. CALA (Thathachar and Sastry, 2004) are learning automata developed for problems with continuous action spaces. CALA are essentially function optimizers: for every action a from their continuous, one-dimensional action space $\mathbb{A}$, they receive a feedback $\beta(x)$ – the goal is to optimize this feedback. CALA have a proven convergence to (local) optima, given that the feedback function $\beta(x)$ is sufficiently smooth. The advantage of CALA over other reinforcement techniques is that it is not necessary to discretize continuous action spaces (actions are simply real numbers). Moreover, they are less complicated to implement and analyze than various other reinforcement techniques for continuous action spaces (Tuyls and Nowe, 2005; Verbeeck *et al*., 2007).

Essentially, CALA maintain a Gaussian distribution from which actions are pulled. In contrast to standard learning automata, CALA require feedback on *two* actions, being the action corresponding to the mean μ of the Gaussian distribution, and the action corresponding to a sample x, taken from this distribution. These actions lead to a feedback $\beta(\mu)$ and $\beta(x)$, respectively, and in turn, this feedback is used to update the probability distribution's μ and σ. More precisely, the update formula for CALA can be written as:

$$\mu = \mu + \lambda \frac{\beta(x) - \beta(\mu)}{\Phi(\sigma)} \frac{x - \mu}{\Phi(\sigma)} \tag{8}$$

$$\sigma = \sigma + \lambda \frac{\beta(x) - \beta(\mu)}{\Phi(\sigma)} \left[\left(\frac{x - \mu}{\Phi(\sigma)} \right)^2 - 1 \right] - \lambda K(\sigma - \sigma_L) \tag{9}$$

In this equation, λ represents the learning rate, set to 0.05 in our case; K represents a large constant driving down σ, which in our case is set to 0.1. The variance σ is artificially kept above a threshold σ_L (set to 10^{-5} in our case), to keep calculations tractable even in case of (near-) convergence. This is implemented using the function:

$$\Phi(\sigma) = \max(\sigma, \sigma_L) \tag{10}$$

The intuition behind the update formula is quite straightforward. First, if the signs of $\beta(x) - \beta(\mu)$ and $x - \mu$ match, μ is increased, otherwise it is decreased. This makes sense, given a sufficiently smooth feedback function: for instance, if $x > \mu$ but $\beta(x) < \beta(\mu)$, we can expect that the optimum is located below the current μ. Second, the variance is adapted depending on how far x is from μ. The term $(\frac{x-\mu}{\Phi(\sigma)})^2 - 1$ becomes positive iff x is more than a standard deviation away from μ. In this case, if x is a better action than μ, σ is increased to make the automaton more explorative. Otherwise, σ is decreased to decrease the probability that the automaton will select x again. If x is not more than a standard deviation away from μ, this behavior is reversed: a 'bad' action x close to μ indicates that the automaton might need to explore more, whereas a 'good' action x close to μ indicates that the optimum might be near. Using this update function, CALA rather quickly converge to a (local) optimum. With multiple (e.g., n) learning automata, every automaton i receives feedback with respect to the joint actions, respectively $\beta_i(\bar{\mu})$ and $\beta_i(\bar{x})$. In this case, there still is convergence to a (local) optimum.

Homo Egualis. As mentioned above, we aim at creating agents that can learn 'human' behavior. The Homo Egualis utility function is a satisfactory model of human behavior in various games, including the games we are letting our agents play, i.e., the Ultimatum Game and the Nash Bargaining Game. Therefore, we use this utility function in our learning agents. More precisely, we use a four-step process. First, every agent i is equiped with a CALA. This automaton selects its actions μ_i and x_i, indicating how much payoff the agent requests. Second, the environment evaluates the joint actions $\bar{\mu}$ and $\bar{x}$ and gives feedback $\beta_i(\bar{\mu})$ and $\beta_i(\bar{x})$, using the rules of the game at hand. In the Ultimatum Game, every agent receives what it has asked for, unless there is not enough reward remaining due to the actions of preceding agents. In this case, the agent

receives what is remaining. In the Nash Bargaining Game, everyone receives what they have asked for, unless the sum of their requests exceeds R. In that case, everyone receives 0. Third, the environment's feedback is mapped to utility values $u_i(\bar{\mu})$ and $u_i(\bar{x})$, using the Homo Egualis function, possibly including punishment (i.e., if any agent experiences a negative utility, all utilities are set to 0). Finally, the utility values are reported to the learning automata, which subsequently update their strategies. Note that the n-player Ultimatum Game requires $n - 1$ automata, whereas the n-player Nash Bargaining Game requires n automata. In the Ultimatum Game, the last agent's behavior is static: he simply rejects if his utility drops below 0. In the Nash Bargaining Game, the last agent is exactly the same as the other agents.

We use the same parameters for the Homo Egualis function (i.e., α_i and β_i) for all agents participating. This makes the analysis and verification of outcomes easier, especially with many agents. Results obtained by giving each agent i private α_i- and β_i-values will be highly similar to our results, but calculating an expected or optimal solution to compare these results with, is more difficult and requires various constraints on the parameters, as we demonstrated in Section 3.1.2.

Extensions to the learning rule. CALA have a proven convergence to a local optimum in the case of smooth and continuous feedback functions. However, as is clearly visible in Figure 3, the feedback function we use (i.e., Homo Egualis) displays a discontinuity: maximum feedback is obtained at a certain value, after which the feedback immediately drops to 0 (if punishment is possible). This leads to two problems, both of which need to be addressed without affecting the convergence of CALA.

The first problem arises when the automaton is near the optimum, and either its x-action or its μ-action is slightly too high. As can be seen from Figure 3, one of the actions will then receive (almost) optimal feedback, whereas the other action receives a feedback of 0. Due to the CALA update function, the μ of the underlying Gaussian will therefore shift drastically (e.g., we observed values of -10^6). As this is a highly undesirable effect, we chose to limit the terms of the update function. More precisely, we limit the term $\beta(x) - \beta(\mu)$ to the interval $[-\Phi(\sigma),\Phi(\sigma)]$. In essence, this addition has the same effect as a variable learning rate, which is not uncommon in literature (e.g., Bowling and Veloso, 2002). In normal cases, i.e., when the automaton is not near the discontinuity, the limit is hardly, if ever, exceeded. Near the discontinuity, it prevents drastic shifts. This addition to the learning rule therefore does not affect convergence.

The second problem arises when both the μ-action and the x-action of the automaton yield a feedback of 0 – i.e., the automaton receives no useful feedback at all. In this case, due to the CALA update function, the underlying Gaussian's μ and σ are not changed. Therefore, in the next learning round, there is a high probability that the automaton again receives a feedback of 0 for both actions. In other words, if this happens, the automaton is very likely to get stuck. We address this issue by including the knowledge that, if both μ and x yield a feedback of 0, the lowest action was nonetheless the best one. Therefore, we set $\beta(x) = \max(\beta(x), \mu - x)$, essentially driving the automaton's μ downward. Once again, in normal cases, the update function remains unchanged. In cases where the automaton receives no useful feedback, it can still update the parameters of the underlying Gaussian. Once again, this addition does not hinder convergence.

4.1.2 Experiments and results

We performed a set of experiments, which are summarized in Table 1. The agents used CALA for learning and the Homo Egualis utility function was applied to the feedback from the environment. The CALA parameters were set as outlined above. The agents started from an initial solution of equal sharing, i.e., all n agents received the same payoff, $R \cdot \frac{1}{n}$. In the experiments, we used $R = 100$. All experiments lasted for 10 000 rounds, were repeated 1 000 times, and the Homo Egualis parameters were set to $\alpha = 0.6$ and $\beta = 0.3$.[7] The number of agents varied between 2, 3, 4, 10 and 100, as denoted under 'Ags.'. Whether or not punishment could be used by the agents is indicated in the 'Pun.' column (i.e., with punishment enabled, agents could reject a solution for which they obtained a negative utility). In the 'Sol.' column, we specify the analytically

[7] In the 'Game' column, we indicate experiments where different settings were used (1. $\beta = 0.7$, 2. repeated 200 times, and 3. repeated 200 times and λ and σ_L lowered by a factor 10).

Table 1 Results of our experiments in the Ultimatum Game and Nash Bargaining Game

Games with 2 players								
Game	Ags.	Pun.	Sol.	Ext. LR.	Avg.	Stdev.	Maint.	Res.
UG[1]	2	no	50.0/50.0	no	50.1/49.9	0.2/0.2	100%	+
UG	2	no	100.0/0.0	no	100.0/0.0	0.0/0.0	100%	+
UG	2	yes	72.7/27.2	no	72.3/27.7	5.5/5.5	100%	+
NBG	2	yes/no	all≥27.2	no	46.5/46.6	2.9/2.7	0%	−
NBG	2	yes/no	all≥27.2	yes	48.2/48.2	2.4/2.4	100%	+
Games with 3 to 10 players								
Game	Ags.	Pun.	Sol.	Ext. LR.	Avg.	Stdev.	Maint.	Res.
UG	3	yes	all>15.8	yes	41.0/41.0/18.0	1.6/1.5/1.7	100%	+
UG	4	yes	all≥11.1	yes	29.0/29.0/29.0/13.0	1.5/1.5/1.5/1.6	100%	+
UG	10	yes	all≥4.0	yes	10.5/10.5/.../6.7	1.1/1.1/.../2.0	100%	+
NBG	3	yes	all≥15.8	yes	33.2/33.1/33.3	1.7/1.7/1.7	100%	+
NBG	4	yes	all≥11.1	yes	24.5/24.5/24.5/24.5	1.6/1.6/1.6/1.6	100%	+
NBG	10	yes	all≥4.0	yes	9.8/9.8/.../9.8	1.2/1.2/.../1.2	100%	+
NBG	3	no	any	yes	33.2/33.1/33.1	1.9/1.9/1.9	93%	∘
NBG	4	no	any	yes	25.0/25.0/25.0/25.0	1.1/1.1/1.1/1.1	93%	∘
NBG	10	no	any	yes	10.0/10.0/.../10.0	0.9/0.9/.../0.9	100%	+
Games with 100 players								
Game	Ags.	Pun.	Sol.	Ext. LR.	Avg.	Stdev.	Maint.	Res.
UG[2]	100	yes	all≥0.4	yes	0.98/0.98/.../4.4	0.4/0.4/.../3.0	81%	∘
UG[3]	100	yes	all≥0.4	yes	0.98/0.98/.../2.2	0.3/0.4/.../1.7	100%	+
NBG[2]	100	yes	all≥0.4	yes	0.94/0.93/.../1.0	0.4/0.4/.../0.4	45%	∘
NBG[3]	100	yes	all≥0.4	yes	0.96/0.92/.../1.0	0.3/0.4/.../0.4	100%	+
NBG[2]	100	no	any	yes	0.89/0.92/.../0.9	0.3/0.3/.../0.3	25%	∘
NBG[3]	100	no	any	yes	0.92/0.96/.../0.9	0.3/0.3/.../0.3	100%	+

determined solution, if there is any, and the constraints that a solution must satisfy if there is no exact solution. This depends on the previous columns. Next, whether or not the extended learning rule was used in the CALA, is indicated under 'Ext. LR.'. Then, we show experimental results (average payoff per agent and standard deviation; the values are separated per agent by a '/'). In every case, we also measure how many times a valid solution was found and subsequently maintained over the full 10 000 rounds (results are displayed under 'Maint.'), i.e., agents all maintain a positive utility. Finally, we indicate the overall results ('Res.') of every set of experiments, i.e., whether they can be considered a success. More precisely, we consider the experiments to be successful (+) if a valid solution was found and maintained in all experiments. The set of experiments is a failure (−) if the agents never find a valid solution. Otherwise, i.e., a solution is found but not always maintained, the set of experiments is neither a success, nor a failure (∘).

Summarizing the results displayed in the table, we may conclude the following. (1) Behavior in the two-agent Ultimatum Game conforms to calculated (and thus human) behavior without requiring the extended learning rule, i.e., the parameter β makes the difference for the first agent between voluntarily sharing 50% and wanting to keep all, and the parameter α can be used to punish overly greedy first agents. (2) The two-agent Nash Bargaining Game does require the extended learning rule for stable solutions to emerge. The solutions in question are compatible with solutions observed in humans. (3) Using Ultimatum and Nash Bargaining Games with 3 to 10 agents, it is still possible to learn calculated (human) solutions. A notable detail is that the first agent in the Ultimatum Game does not exploit the other agents; instead, only the last agent is exploited. This is due to the fact that all agents except the last one are learning simultaneously. An example learning episode is illustrated in Figure 6. There is currently no 'human' data to verify

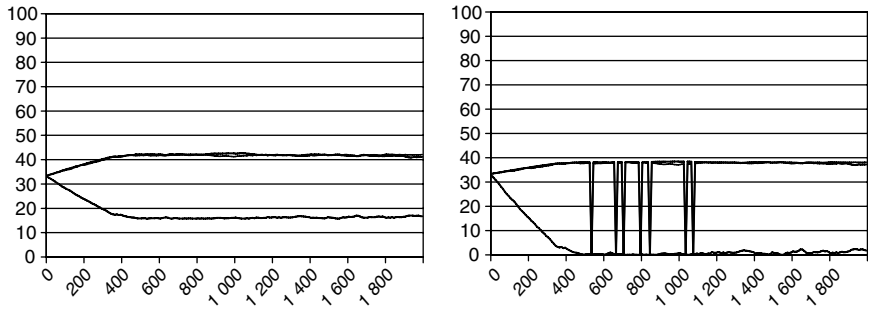

Figure 6 Three agents playing the Ultimatum Game. The automata learn for 2 000 iterations. The figure shows traces of the agents' payoffs (left) and utility values (right). The last agent's utility is quickly minimized to a value of 0, corresponding to a payoff of ~ 18. Punishment by the last agent ensures that the other two agents do not take even more of the reward (i.e., the vertical lines in the utility trace). The other two agents converge to an equal split of the remaining reward (i.e., 82).

this behavior, but it is known from other experiments that humans indeed tend to share, even with high stakes (see Figure 5). (4) Using 100 agents, the parameters used for the CALA had to be lowered to make learning possible, simply because agents should learn to obtain a payoff of only 1, which is hard if the learning mechanism is overly noisy. (5) Disabling the possibility to punish in the Nash Bargaining Game makes the game less easy to solve; valid solutions are not always found and maintained. Due to punishment, agents become slightly more conservative.

As an overall conclusion, we see that the architecture discussed, coupling CALA and the Homo Egualis utility function, is indeed able to find and maintain solutions to the Ultimatum Game and the Nash Bargaining Game. Moreover, these solutions indeed correspond to solutions found by humans, as reported in literature. Note that this correspondance is more qualitative than quantitative. Humans behave rather differently on a quantitative level (e.g., there is a broad variation in the minimal payoff accepted by humans in the Ultimatum Game), but follow the same tendencies on a qualitative level (i.e., hardly anyone agrees with the individually rational solution of accepting 1 cent). The parameters of the Homo Egualis utility function (i.e., α and β) allow for a broad range of 'human-like' behavior. Obviously, it remains an interesting challenge to set these parameters to adequate (quantitative) values when interacting with real humans. A possible way of dealing with this challenge is to estimate the parameters based on a few 'test' games with these humans, as indicated by Fehr and Schmidt (1999) and Dannenberg *et al.* (2007).

4.2 Toward priority awareness

Note that the methodology described above can also be applied to priority-aware agents. After all, priority awareness is basically an extension to inequity aversion: instead of perceiving their actual payoffs, agents perceive a payoff that is influenced by their priority. Thus, convergence results remain the same. The only difference here lies in the verification of obtained solutions. Since agents now use four instead of two parameters, calculating expected solutions becomes very difficult. We addressed a similar problem with the two inequity-aversion parameters α_i and β_i, by simply setting them to the same values for all agents. With priority awareness, setting all parameters to the same values for all agents would not lead to any new results, since priority awareness and inequity aversion are identical in this case. Thus, even though it is very possible to learn priority-aware solutions, analytically verifying these solutions becomes substantially more difficult.

4.3 Reciprocal fairness in actual multi-agent settings

As we explained in Section 3.3, many researchers attempted to come up with mechanisms that could explain cooperation in social dilemmas. Many mechanisms were verified by engineering them into actual adaptive multi-agent systems and studying the dynamics of these systems in

various social dilemmas, such as the Prisoners' Dilemma, the Ultimatum Game (see above), and the Public Goods Game. If these systems are subsequently shown to display behavior that is in line with human fair behavior, we can state that the underlying mechanisms can be used as descriptive as well as prescriptive models of human fairness. Various such mechanisms have been coined and experimented upon, including volunteering (which has been addressed in Section 3.3), reputation systems and models of local interaction. We will provide a brief overview here.

Reputation. Various researchers (e.g., Nowak *et al*., 2000; Sigmund *et al*., 2001; Panchanathan and Boyd, 2004; Fehr, 2004; Falk and Fischbacher, 2006)] argue that fairness (or, alternatively, altruistic punishment) is not possible without a representation of reputation. To support this claim, the behavior of agents driven by such systems is analyzed, mostly from the perspective of evolutionary game theory (Gintis, 2001). In Nowak *et al*. (2000) for instance, the Ultimatum Game is converted to a 'mini-game' with only two strategies per agent (offering a high or a low payoff, as well as wishing to obtain such offers). The Nash equilibrium of this game is to offer and accept a low payoff. In contrast, the fair solution is to offer and accept a high payoff. Using replicator equations, which describe population dynamics where successful strategies spread, it is derived that indeed, a population of agents playing this game will converge to the Nash equilibrium. Next, the possibility is introduced that agents can obtain information about previous encounters, i.e. previously accepted offers. In this case, depending on initial conditions, the population either converges to the Nash equilibrium, or to the fair solution. In the real Ultimatum Game, i.e., with continuous strategies instead of only two, the same happens. The more games are played and the more reputation spreads, the faster the system converges to fair solutions. This, as Nowak *et al*. (2000) note, agrees well with other findings on the emergence of cooperation or of bargaining behavior. Similarly, Sigmund *et al*. (2001) show that reputation is significantly more effective in combination with a mechanism that allows punishing those who have a bad reputation, than with rewarding those who have a good reputation.

Image scoring and good standing. Image scoring (Nowak and Sigmund, 1998, 2005; Wedekind and Milinski, 2000) is a practical implementation of the idea of maintaining a reputation value for all individuals in a certain population. Strategies based on image scoring are of the kind where one gives help only to those whose score is above a certain threshold (Leimar and Hammerstein, 2001). In practice, an individual's score increases on every occasion he donates aid to a recipient and decreases when there is an opportunity to help someone in need but no help is offered. Analysis and computer simulations show that image scoring may indeed lead to cooperation and reciprocal altruism. In addition, Lotem *et al*. (1999) show that cooperation may actually be increased by adding a small fraction of agents to the population that is physically unable to display cooperative behavior. Whether image scoring indeed forms a satisfactory implementation of reputation, is debatable. For instance, Leimar and Hammerstein (2001) argue that analytical arguments show it would not be in an individual's interest to use image-scoring strategies. Image scoring only works in case of strong genetic drift or a very small cost of actually being altruistic. As an alternative, they propose the strategy of aiming for 'good standing', which is inspired by Sugden (1986) and demonstrated to be (potentially) evolutionarily stable and dominant over image scoring. In contrast to the image-scoring model, the good-standing model initially attributes a high value to all individuals. Individuals may lose standing if they refuse to help others that have good standing.

Local interactions. In addition to studies being performed on internal mechanisms of agents, such as their perception of reputation, intentions or voluntary participation, there are also studies focusing on external factors that may lead to fairness. Most notably, researchers argue that (1) humans do not interact on a random basis, as traditionally assumed by population dynamics; instead, human interactions, like many other natural phenomena, seem to be organized in scale-free or small-world networks (Santos and Pacheco, 2005; Santos *et al*., 2006a, b). Moreover,

	a_{21}	a_{22}
a_{11}	$(1, 1)$	(s, t)
a_{12}	(t, s)	$(0, 0)$

Figure 7 A generalized social dilemma, as presented by Santos *et al.* (2006b). Various dilemmas can be constructed using two variables $s \in [-1, 1]$ (sucker's payoff) and $t \in [0, 2]$ (temptation).

(2) humans are able to adjust their social ties: in case they interact with a person they turn out not to like, they may refuse to interact with this person again; in this case, for instance, a cooperative player is not forced to become defective once he interacts with a defector; instead, the player can choose to stay cooperative, but just not with this defector (Santos *et al.*, 2006c). We elaborate on these two arguments. In Santos and Pacheco (2005) and Santos *et al.* (2006a, b), the effect of local interactions on cooperation is studied using a generalized class of social dilemmas, characterized by a payoff matrix for both players that is parametrized using two variables $s \in [-1, 1]$ (sucker's payoff) and $t \in [0, 2]$ (temptation), see Figure 7. Many values of s and t, i.e., many social dilemmas, are played using adaptive agents. Agents may imitate the strategies of other agents they played against, with a probability similar to the difference in payoff (i.e., a successful agent is more likely to be imitated). The agents are paired based on various network topologies, including single-scale networks and scale-free networks, as can be generated with the well-known algorithm by Barabasi and Albert (1999). It is shown that cooperation is the dominant strategy in scale-free networks for many more values of s and t than in other networks; thus, the heterogeneity of the network with which agents interact, increases the probability of cooperation becoming the dominant strategy. More precisely, cooperation prevails once there is a majority of cooperators, the hubs (densely connected agents) in the network are interconnected and the network is sparse. However, in any other case, cooperation is (still) doomed. Santos *et al.* (2006c), therefore continue their studies and argue that we should not assume that connections in social networks are static; after an unsatisfactory interaction, agents may choose to rewire their connection to someone else. Basically, everyone would like to connect to a cooperator (since this leads to a higher payoff both as a defector or as a cooperator). The probability that such rewiring is allowed, is chosen to depend on the fitness (average payoff) of the agent wanting to be rewired. The authors show that, with a sufficiently high rewiring probability, full cooperation can be reached in all social dilemma games experimented on. The resulting networks are realistic, with a high average connectivity and associated single- to broad-scale heterogeneity. Note that this idea is closely related to the concept of volunteering, as introduced by, e.g., Hauert *et al.* (2002, 2007); agents can choose whether or not to interact with certain other agents.

4.4 Other computational models for multi-agent systems

In this section, we give an overview of three other fields closely related to achieving desired outcomes in a collective of agents, i.e. mechanism design, computational social choice, and normative systems. In these fields, different ways of achieving certain outcomes are studied. Outcomes may vary from individually rational outcomes to 'fair' outcomes, e.g., outcomes in which a certain social utility is maximized.

4.4.1 Mechanism design

Economy-based collaboration mechanisms are very popular in multi-agent research. Mechanism design (Parkes, 2008) studies the art of designing the rules of a game such that a specific outcome is achieved. Most of the results in mechanism design have been established by economists, but some mathematicians, computer scientists and electrical engineers also work in the field.

As in the research outlined in this paper, mechanism design assumes that players' individually rational actions may not lead to a desired global outcome. Thus, designers set up a structure in which each player has an incentive to behave as intended. They commonly try to achieve truthfulness, individual rationality, budget balance, and social welfare. More advanced mechanisms attempt to resist harmful coalitions of players. For instance, in a multi-agent systems context, the utility of the system as a whole can be determined using a social-choice function, such as

maximising the total utility gained across all agents (see below). For a comprehensive overview, see Chevaleyre *et al*. (2006) and Jackson (2000).

From a computational point of view, some specific issues arise in mechanism design. To start with, unlike with game theoretic assumptions, software agents do not possess unbounded computational power to calculate equilibrium strategies. Theory focusses on centralized mechanisms, but the infrastructure might be unable to compute the outcome because the problem might be intractable. Furthermore, communication between the agents is not necessarily cost- or error-free and the system might be dynamic, with agents entering or leaving the system over time. Current state-of-the-art research in computational mechanism design addresses one or more of these added computational issues (Parkes, 2008).

Previous work of two of the authors of this paper was situated in the area of fairness and mechanism design. More precisely, in Verbeeck *et al*. (2007), agents are motivated by the Homo Egualis fairness model to learn how to fairly share public resources in single stage conflicting interest games. However, rather than using the Homo Egualis utility function directly, the authors developed a coordinated exploration mechanism in which agents learn to select sub-optimal actions for some period of time, in favor of less performing agents.

4.4.2 Computational social choice

Another area in which fairness is extensively studied in a computational manner is that of computational social choice (see Chevaleyre *et al*. (2006, 2007) for a comprehensive overview). This area encompasses many interesting problems at the interface of social choice theory and computer science, for instance fair division in resource allocation (Chevaleyre *et al*., 2006). In this case, fairness conditions and mechanisms relate to the well-being of society as a whole. This well-being can be measured in various ways, such as utilitarian social welfare (i.e., maximized average payoff), egalitarian social welfare (i.e., maximized minimal payoff), or Pareto-optimality. We argue that another measure for the well-being of society as a whole should be introduced, i.e., a definition of fairness that is backed up by numerous well-documented experiments with human subjects.

4.4.3 Institutional and social norms

As a final contribution in this section, we mention research in the area of norms and institutions (Rodriguez-Aguilar, 2003; Vazquez-Salceda *et al*., 2005; Aldewereld, 2007). There are interesting parallels between this research and the research described in the previous sections.

Norms can be used as an environment-driven coordination mechanism, and are an alternative to classical, agent-centered coordination. Especially in open multi-agent systems, i.e., systems in which agents may be heterogeneous or designed by different parties, one cannot assume that all agents pursue the same goal or have the same internal procedures. In fact, the goal of certain agents may be to disrupt the system or to exploit the other agents in the system. In situations such as these, norms may help, since they allow agents to make predictions about others and to direct their own actions toward desirable behavior.

Enforcing norms is a complex problem, since norms are usually represented in formalisms that have a declarative nature, but should be translated to an operational implementation (Vazquez-Salceda *et al*., 2005). In other words, it is only known how and when a norm is violated, but many questions, such as how violations should be detected and which responsive actions (punishments) are appropriate, remain. Note that this distinction between operational and declarative representations of norms has some parallels in the distinction between descriptive and prescriptive (or computational) models of fairness (although obviously, norms can describe much more than just considerations for fairness).

In natural societies, norms and associated punishments emerge over time, either spontaneously or deliberately. Societies use social constraints (norms) to regulate relations among their members, such as customs, traditions, regulations or laws. A collection of such constraints is called an *institution* (North, 1990; Scott, 2001). In the context of multi-agent systems, such an institution may be established for reasons of *trust*. For instance, on an online auction site, the institutional norm may

be that the winning bidder is obliged to purchase the item in question, and the item is only delivered after the bidder transferred his payment to the other party. People that do not adhere to this norm may be punished, i.e., they may obtain a bad reputation, which is displayed on the auction site. As a result, others may be unwilling to interact with them. With such a norm and associated 'punishment' in place, both buyers and sellers know what to expect, i.e., a form of trust emerges on the simple basis that all agents involved (have to) agree to follow certain institutional norms.

In addition to being institutional (i.e., enforced by the environment), norms may also be appointed between individual agents; in this case, they are referred to as social norms. Fairness may be reflected in both types of norms. Studying the emergence of social norms in agent systems is recognized as an important research track, since it may improve coordination and functioning of the agent system (Sen and Airiau, 2007). However, making social norms emerge is difficult, since the 'initiative' should not be with individual agents too much—otherwise, agents may still choose to adopt their own incompatible norms. Another question of interest for research on fairness in multi-agent systems, is whether the emerged agent norms are aligned with human norms.

If we define multi-agent systems as systems that may include humans in addition to autonomous software or robots, it becomes impossible to predefine or even restrict the internal procedures of all the agents in such systems. Some kind of norm enforcement is therefore necessary to prevent agents from achieving undesirable goals. For instance, similar to the auction example, agents that do not adhere to norms may obtain a bad reputation in the system. In fact, most work described previously (e.g., using the Homo Egualis utility function, introducing reputation, volunteering, etc.) can be explained in the context of norms and institutions. The norms ensure that agents can trust each other and can understand (or learn) how to achieve their goals within the institution. Developing institutional norms that drive agents to a desired solution, is a problem that is closely related to mechanism design (see above). Although current research in the area of fairness in multi-agent systems is mostly limited to systems in which all agents are designed by the same party, many ideas would also be applicable in open systems.

5 Conclusion

This paper presents an overview of work in the area of fairness for multi-agent systems. In essence, there are two distinct reasons for incorporating fairness concerns. First, multi-agent systems often perform tasks for humans, or even interact with them. Since research shows that humans are not individually rational, the classical, individually rational agent model may not be sufficient to obtain alignment with human expectations. Second, there are multiple examples of multi-agent interactions in which following an individually rational strategy actually leads to bad results. In contrast, strategies that are more aware of concepts such as social welfare or fairness obtain better results.

In Section 2 of this paper, we first gave a brief overview of the classical, individually rational agent model, which is driven by concepts derived from game theory, such as the Nash equilibrium. Next, in Section 3, we discussed the concept of human fairness, which has been described in three different models, i.e., inequity aversion, priority awareness, and reciprocal fairness. Each model was developed to explain the human behavior observed in numerous experiments. The first two models focus on the human response to observed difference in payoff. Essentially, both models describe the human attitude with respect to other individuals doing better or worse. The priority-aware model introduces the notion of a priority, which allows certain individuals to be better than others. The third model, reciprocal fairness, encompasses various explanations for the emergence of cooperation in repeated interactions, most notably reputation and punishment. Section 4 discussed how descriptive models may be translated to computational models, i.e., models that prescribe desired behavior to agents in multi-agent systems. We discussed this for each of the three descriptive models introduced. Moreover, we introduced various other computational models for multi-agent systems, which do not necessarily focus on fairness.

Interdisciplinary research in the area of fairness for multi-agent systems provides many interesting opportunities, both for researchers wishing to understand and describe human behavior, as for researchers trying to make multi-agent systems more aware of such behavior. For instance, a model that integrates the various descriptive models of human fairness into one computational model, may be able to address many more problems than the current computational models, and in addition may reveal more to us about how humans work. Especially, it should be noted that current research is often performed on rather abstract problems, i.e., mostly single-stage games. Fairness is definitely also important in more complex problem domains, such as auctions, planning and scheduling. This line of research requires more advanced and more extensive descriptions of the human way of thinking, which subsequently may be integrated in multi-agent systems. Also, for every problem domain, it should be analyzed whether adding fairness indeed offers a benefit in comparison to a purely rational system. Even performance measures that may facilitate such analysis are currently debatable. For instance, choosing to minimize the average waiting time of resources in multi-agent resource allocation problems leads to significantly different results than choosing to minimize the maximum waiting time or the variance. In conclusion, this paper possibly only presents the first steps on the long road of trying to understand ourselves and trying to get computers to understand us as well.

Acknowledgments

The research reported here is supported by the 'Breedtestrategie' programme of Maastricht University. This paper is partially based on previously published work—we thank our co-authors and the anonymous reviewers for their valuable contributions. Moreover, we thank Huib Aldewereld for a refreshing view on the material presented in Section 4.4.3. In addition, we would like to thank Anton de Vries for his assistance in developing the prioritized-Ultimatum-Game survey, and our respondents for their time. Finally, we acknowledge Lumière Maastricht for providing us with cinema tickets that we could use as an incentive for our respondents.

References

Aldewereld, H. 2007 Autonomy vs. conformity: an institutional perspective on norms and protocols. PhD thesis, Universiteit Utrecht.

Aristotle. Politics, Book II, Chapter III, 1261b; translated by Benjamin Jowett as *The Politics of Aristotle*: Translated into English with Introduction, Marginal Analysis, Essays, Notes and Indices, volume **1 of 2**. Oxford: Clarendon Press, 1885.

Barabasi, A.-L. and Albert, R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512.

Basu, K. 1994. The traveler's dilemma: paradoxes of rationality in game theory. *American Economic Review* **84**(2), 391–395.

Basu, K. 2007 The traveler's dilemma. *Scientific American* **296**(6), 68–73.

Bowles, S., Boyd, R., Fehr, E. and Gintis, H. 1997 Homo reciprocans: a research initiative on the origins, dimensions, and policy implications of reciprocal fairness. *Advances in Complex Systems* **4**, 1–30.

Bowling, M. H. and Veloso, M. M. 2002 Multiagent learning using a variable learning rate. *Artificial Intelligence* **136**, 215–250.

Boyd, R., Gintis, H., Bowles, S. and Richerson, P. 2003 The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3531–3535.

Boyd, R. and Mathew, S. 2007 A narrow road to cooperation. *Science* **316**, 1858–1859.

Charness, G. and Rabin, M. 2002 Understanding social preferences with simple tests. *Quarterly Journal of Economics* **117**, 817–869.

Chevaleyre, Y., Dunne, P. E., Endriss, U., Lang, J., Lemaître, M., Maudet, N., Padget, J., Phelps, S., Rodriguez-Aguilar, J. A. and Sousa, P. 2006 Issues in multiagent resource allocation. *Informatica* **30**, 3–31.

Chevaleyre, Y., Endriss, U., Lang, J. and Maudet, N. 2007 A short introduction to computational social choice. In *Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM-2007)*, vol. 4362 of LNCS, Berlin, Heidelberg, Germany: Springer-Verlag, pp. 51–69.

Dannenberg, A., Riechmann, T., Sturm, B. and Vogt, C. 2007 Inequity aversion and individual behavior in public good games: an experimental investigation. SSRN eLibrary.

Dawkins, R. 1976 *The Selfish Gene*. Oxford: Oxford University Press.

De Jong, S., Tuyls, K. and Sprinkhuizen-Kuyper, I. 2006 Robust and scalable coordination of potential-field driven agents. In *Proceedings of IAWTIC/CIMCA* 2006, Sydney.

De Jong, S., Tuyls, K. and Verbeeck, K. 2008a Artificial agents learning human fairness. In Accepted at the *International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'08)*.

De Jong, S., Tuyls, K., Verbeeck, K. and Roos, N. 2008b Priority awareness: towards a computational model of human fairness for multi-agent systems. *Adaptive Agents and Multi-Agent Systems III—Lecture Notes in Artificial Intelligence*, 4865.

Denning, D. 1982 *Cryptography and Data Security*. Boston, USA: Addison-Wesley.

Erev, I. and Roth, A. E. 1998 Predicting how people play games with unique, mixed strategy equilibria. *American Economic Review* **88**, 848–881.

Falk, A. and Fischbacher, U. 2006 A theory of reciprocity. *Games and Economic Behavior* **54**, 293–315.

Fehr, E. 2004 Don't lose your reputation. *Nature* **432**, 499–500.

Fehr, E. and Gaechter, S. 2000 Fairness and retaliation: the economics of reciprocity. *Journal of Economic Perspectives* **14**, 159–181.

Fehr, E. and Gaechter, S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140.

Fehr, E. and Schmidt, K. 1999 A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* **114**, 817–868.

Ferber, J. 1999 *Multi-Agent Systems. An Introduction to Distributed Artificial Intelligence*. Boston, USA: Addison-Wesley.

Gintis, H. 2001 *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton, USA: Princeton University Press.

Gueth, W., Schmittberger, R. and Schwarze, B. 1982 An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior and Organization* **3**(4), 367–388.

Hamilton, W. 1964 The genetical evolution of social behaviour I and II. *Journal of Theoretical Biology* **7**, 1–16, 17–52.

Hardin, G. 1968 The tragedy of the commons. *Science* **162**, 1243–1248.

Hauert, C., Monte, S. D., Hofbauer, J. and Sigmund, K. 2002 Volunteering as red queen mechanism for cooperation in public goods games. *Science* **296**, 1129–1132.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. and Sigmund, K. 2007 Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907.

Henrich, J. 2004 Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization* **53**, 3–35.

Henrich, J. and Boyd, R. 2001 Why people punish defectors. *Journal of Theoretical Biology* **208**, 79–89.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E. and Gintis, H. 2004 *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.

Henrich, J., McElreath, R., Barr, A., Ensimger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N. *et al*. 2006 Costly punishment across human societies. *Science* **312**, 1767–1770.

Jackson, M. 2000 Mechanism Theory. *Humanities and Social Sciences* October, 228–277.

Jennings, N., Sycara, K. and Wooldridge., M. 1998 A roadmap of agent research and development. *Autonomous agents and Multi-Agent Systems* **1**, 275–306.

Kalagnanam, J. and Parkes, D.C. 2004 Auctions, bidding and exchange design. In Simchi-Levi, D., Wu, S. D. and Shen, M. (eds.), *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era*, International Series in Operations Research and Management Science, ch. 5. Dordrecht, Netherlands: Kluwer Academic Publishers, pp. 1–84.

Kalisch, G., Milnor, J., Nash, J. and Nering, E. 1952 Some experimental n-person games. Technical report, The Rand Corporation U.S. Air Force.

Keller, L. and Ross, K. G. 1998 Selfish genes: a green beard in the red fire ant. *Nature* **394**(6693), 573–575.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V. and Fehr, E. 2006 Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* **314**, 829–832.

Leimar, O. and Hammerstein, P. 2001 Evolution of Cooperation through Indirect Reciprocity. *Proceedings: Biological Sciences* **268**(1468), 745–753.

Lotem, A., Fishman, M. and Stone, L. 1999 Evolution of Cooperation between individuals. *Nature* **400**, 226–227.

Mao, X., ter Mors, A., Roos, N. and Witteveen, C. Agent-based scheduling for aircraft deicing. In Schobbens, P.-Y., Vanhoof, W. and Schwanen, G. (eds.) *Proceedings of the 18th Belgium–Netherlands Conference on Artificial Intelligence*, BNVKI, October 2006. pp. 229–236.

Maynard-Smith, J. 1982 Evolution and the Theory of Games. Cambridge University Press.

Milinski, M., Semmann, D. and Krambeck, H. 2002 Reputation helps solve the tragedy of the commons. *Nature* **415**, 424–426.

Narendra, K. and Thathachar, M. 1989 *Learning Automata: An introduction*. Boston, USA: Prentice-Hall International.

Nash, J. 1950a Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the United States of America* **36**, 48–49.

Nash, J. 1950b The bargaining problem. *Econometrica* **18**, 155–162.

North, D. 1990 *Institutions, Institutional Change and Economic Performance*. 1st edn., Cambridge: Cambridge University Press.

Nowak, M., Page, K. and Sigmund, K. 2000 Fairness versus reason in the Ultimatum Game. *Science* **289**, 1773–1775.

Nowak, M., Sasaki, A., Taylor, C. and Fudenberg, D. 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650.

Nowak, M. A. and Sigmund, K. 1998 Evolution of indirect reciprocity by image scoring. *Nature* **393**(6685), 573–577.

Nowak, M. A. and Sigmund, K. 2005 Evolution of indirect reciprocity. *Nature* **437**(7063), 1291–1298.

Nydegger, R. and Owen, H. 1974 Two-person bargaining, an experimental test of the Nash axioms. *International Journal of Game Theory* **3**, 239–250.

Oosterbeek, H., Sloof, R. and van de Kuilen, G. 2004 Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Experimental Economics* **7**, 171–188.

Osborne, J. and Rubinstein, A. 1994 *A Course in Game Theory*. Boston, USA: MIT Press.

Panchanathan, K. and Boyd, R. 2004 Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* **432**, 499–502.

Parkes, D. C. 2008 Computational Mechanism Design. In *Lecture notes of Tutorials at 10th Conference. on Theoretical Aspects of Rationality and Knowledge (TARK-05)*. Institute of Mathematical Sciences, University of Singapore.

Preist, C. and van Tol, M. 1998 Adaptive agents in a persistent shout double auction. In *ICE '98: Proceedings of the First International Conference on Information and Computation Economies*, New York: ACM Press, pp. 11–18.

Rabin, M. 1993 Incorporating Fairness into Game Theory and Economics. *American Economic Review* **83**, 1281–1302.

Rodriguez-Aguilar, J. A. 2003 On the design and construction of agent-mediated electronic institutions. PhD thesis, Monografies de l'Institut d'Investigació en Intelligència Artificial.

Roth, A. and Malouf, M. 1979 Game theoretic models and the role of information in bargaining. *Psychological Review* **86**, 574–594.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E. and Cohen, J. D. 2003 The neural basis of economic decision-making in the ultimatum game. *Science* **300**, 1755–1758.

Santos, F. and Pacheco, J. 2005 Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters* **95**, 98–104.

Santos, F., Pacheco, J. and Lenaerts, T. 2006a Emergence of cooperation in heterogeneous structured populations. *Proceedings of the 10th International Conference on the Simulation and Synthesis of Living Systems (Alife X)*. 432–437.

Santos, F., Pacheco, J. and Lenaerts, T. 2006b Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3490–3494.

Santos, F., Pacheco, J. and Lenaerts, T. 2006c Cooperation prevails when individuals adjust their social ties. *PLoS Computational Biology* **2**(10), 1284–1291.

Scott, W. 2001 *Institutions and Organizations*, 2 edn. Thousand Oaks, USA: Sage.

Sen, S. and Airiau, S. 2007 Emergence of norms through social learning. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence* pp. 1507–1512.

Shoham, Y., Powers, R. and Grenager, T. 2007 If multi-agent learning is the answer, what is the question?. *Artificial Intelligence* **171**(7), 365–377.

Sigmund, K., Hauert, C. and Nowak, M. 2001 Reward and punishment. *Proceedings of the National Academy of Sciences of the United States of America* **98**(19), 10757–10762.

Simon, H. 1957 *Models of Man*. London, UK: John Wiley.

Simon, H. 1972 Theories of Bounded Rationality. *Decision and Organization* **1**, 161–176.

Simon, H. and Newell, A. 1972 *Human Problem Solving*. Englewood Cliffs, USA: Prentice Hall.

Sugden, R. 1986 *The Economics of Rights, Co-operation and Welfare*. Oxford: Basil Blackwell.

Thathachar, M. and Sastry, P. 2004 *Networks of Learning Automata: Techniques for Online Stochastic Optimization*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Tuyls, K. and Nowe, A. 2005 Evolutionary game theory and multi-agent reinforcement learning. *The Knowledge Engineering Review* **20**(01), 63–90.

Van Huyck, J., Battalio, R., Mathur, S. and Ortmann, A. 1995 On the origin of convention: evidence from symmetric bargaining games. *International Journal of Game Theory* **34**, 187–212.

Vaughn, K. L. 1987 *The New Palgrave: A Dictionary of Economics*, vol. 2. ch. Invisible Hand. London, UK: Macmillan.

Vazquez-Salceda, J., Aldewereld, H. and Dignum, F. 2005 Norms in multiagent systems: from theory to practice. *Computer Systems Science and Engineering* **20**(4), 225–236.

Verbeeck, K., Nowé, A., Parent, J. and Tuyls, K. 2007 Exploring selfish reinforcement learning in repeated games with stochastic rewards. *Journal of Autonomous Agents and Multi-Agent Systems* **14**, 239–269.

von Neumann, J. and Morgenstern, O. 1944 *Theory of Games and Economic Behaviour*. Princeton, USA: Princeton University Press.

Wedekind, C. and Milinski, M. Cooperation through image scoring in humans. *Science* **288**(5467), 850–852.

Weiss, G. 1999 *Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence*. Boston, USA: MIT Press.

Weyns, D., Boucke, N., Holvoet, T. and Schols, W. 2005 Gradient field-based task assignment in an AGV transportation system. In *Proceedings of EUMAS* pp. 447–458.

Yaari, M. and Bar-Hillel, M. 1981 On dividing justly. *Social Choice and Welfare* **1**, 1–24.

Yamagishi, T. 1986 The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* **51**(1), 110–116.

Zollman, K. J. 2008 Explaining fairness in complex environments. *Politics, Philosophy, and Economics* **7**(1), 81–97.