Priority Awareness: Towards a Computational Model of Human Fairness for Multi-agent Systems

Conference Paper · January 2007

DOI: 10.1007/978-3-540-77949-0_9 · Source: DBLP

4 authors:

Steven de Jong
TNO

Karl Tuyls
DeepMind and University of Liverpool

Katja Verbeeck
Odisee

Nico Roos
Maastricht University

Considerations for fairness in multi-agent systems

Steven de Jong, Karl Tuyls, Katja Verbeeck and Nico Roos

MICC/IKAT, Maastricht University, The Netherlands
{steven.dejong,k.tuyls,k.verbeeck,roos}@micc.unimaas.nl

Abstract. Typically, multi-agent systems are designed assuming perfectly rational, self-interested agents, according to the principles of classical game theory. However, research in the field of behavioral economics shows that humans are not purely self-interested; they strongly care about whether their rewards are *fair*. Therefore, multi-agent systems that fail to take fairness into account, may not be sufficiently aligned with human expectations. Two important motivations for fairness have already been identified and modelled, being (i) inequity aversion and (ii) reciprocity. We identify a third motivation that has not yet been captured: priority awareness. We show how priorities may be modelled and discuss their relevance for multi-agent research.

1 Introduction

Modelling agents for a multi-agent system requires a thorough understanding of the type and form of interactions with the environment and other agents in the system, including any humans. Since many multi-agent systems are designed to interact with humans or to operate on behalf of them [1, 2], agents' behavior should often be aligned with human expectations. If a multi-agent system is insufficiently aligned, humans may not understand and even reject it.

Usually, multi-agent systems are designed according to the principles of a standard game-theoretical model. More specifically, the agents assume complete knowledge of the environment, are perfectly rational and optimize their individual payoff disregarding what this means for the utility of the entire population. Experiments in behavioral economics have taught us that humans often do *not* behave in such a self-interested manner [3–5]. Instead, they take into account the effects of their actions on others; i.e., they strive for *fair* solutions and expect others to do the same. Therefore, multi-agent systems using only standard game-theoretical principles risk being insufficiently aligned with human expectations.

To avoid this problem, designers of multi-agent systems should take the human conception of fairness into account. If the motivations behind human fairness are sufficiently understood and modelled, the same motivations can be applied in multi-agent systems. This interesting track of research ties in with the descriptive agenda formulated by Shoham [6] and the objectives of evolutionary game theory [5, 7].

In the remainder of this paper, we first discuss related work in the area of fairness models. Then, we look at problems in which priorities play a role. We show that current models do not predict human behavior in such problems. Next, we provide our model, and perform experiments to show that the model performs a much better prediction of human behavior. We conclude with some directions for future work.

2 Related work

Already in the 1950's people started looking at fairness, for instance in the Nash bargaining game [8]. Recently, research in behavioral economics and evolutionary game theory has examined human behavior in the ultimatum game and the public goods game [3–5, 9, 10]. In all cases, it was observed that standard game theoretical models predict a very selfish outcome in comparison to the fair outcomes reached by human players. The current state of the art describes and models two main motivations for human fairness.

The first motivation for human fairness is *inequity aversion*. In [4], this 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"*. To model inequity aversion, an extension of the classical game theoretic actor is introduced, named homo egualis [4, 5]. Agents using the homo egualis utility function care more about inequity if it is to their disadvantage than if it is to their advantage. Research with human subjects provides strong evidence that this is a valid assumption [4].

Second, various researchers argue that fairness arises most notably in the presence of *reciprocity* [3–5, 11, 12], i.e., agents can reward and punish others. To model this motivation for fairness, a second actor is developed, named homo reciprocans [3, 5]. This actor responds friendly to cooperation of others by maintaining this level of cooperation, and responds to defection by retaliating against the offender(s), even if this reduces her own payoff. In iterated games, this usually leads to an equilibrium in which all players behave in a fair way.

3 Fairness in priority problems

3.1 Priority problems

In many situations, humans actually consider it fair that each of them gets a (slightly) different reward, because they take into account additional information. For instance, everybody can agree that priority mail should be delivered faster than regular mail. Such situations are present in many common applications of multi-agent systems and should therefore be addressed when looking at fairness.

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. We will denote this additional information 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 Example and human response

In a certain furniture store, customers wait at a service desk while employees fetch the items ordered. Obviously, a customer will not be happy if he observes that five other customers are helped while he is still waiting at the service desk. Neither will customers requesting the most popular item be pleased when they discover that they all have to wait for five minutes because another customer has ordered an extremely rare item. In

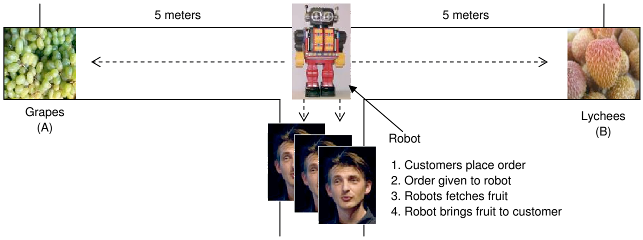

Fig. 1. A human fairness test in a small shop environment.

other words, customers are willing to accept that someone who orders a common item (i.e., an item with high probability of being ordered) is helped slightly more quickly than someone who orders a rare item (with low probability of being ordered).

To investigate how humans deal with such a problem, in [13], we discuss a test of human fairness in a small shop environment, inspired by this example. The store at hand sells two types of fruit (see Figure 1), located at A and B respectively. The human respondents are given the probabilities that customers wish to order the fruit at A (say, p) or the fruit at B ($1 - p$) and are then asked to place the robot somewhere on the line AB, such that all customers will be satisfied. We found that human behavior in this test is not appropriately predicted by analytical measures. More precisely, if $p > 0.5$, analytical measures would place the robot at A (in order to minimize expected waiting time) or in the middle (in order to minimize the variance in waiting time). Humans on the other hand tend to place the robot somewhere between A and the middle. Thus, they perform a trade-off between the expected waiting time and the variance (i.e., they choose a more fair solution). The larger p becomes, the closer to A the robot is placed. We then tried to apply the two known motivations for fairness (inequity aversion and reciprocity) and found that these did not predict human behavior either.

3.3 Why current models do not work

More generally speaking, both inequity-averse and reciprocal models cannot be used to sufficiently explain or predict the human conception of fairness in problems where priorities play a role.

In [4, 5], the homo egualis model, which is inequity-averse, is applied to the ultimatum game and the public goods game. Obtained results are in line with human behavior, as long as the model's parameters are chosen carefully. However, as clearly indicated in [4], in many situations, people actually like to be better than other people, and the homo egualis actor does not model this behavior. In priority problems, people do not only tolerate, but even encourage inequity; agents with a high priority should actually like to be better than agents with a low priority (to a certain degree).

Asking our players to place the robot "such that all customers are satisfied" implicitly introduces a threat of punishment to this player; if customers are dissatisfied, they might not return to the shop in question. The outcomes of the test might therefore be explainable in terms of reciprocity. However, there are two problems when trying to apply a reciprocal model to the customer game. First, reciprocal models assume a setting in which players receive explicit punishments (or rewards) after performing a certain action. In case of our test, no explicit feedback was given to the respondents. Second, the reciprocal model does not provide us with a 'recipe' for fair behavior, neither in standard problems (such as the ultimatum game), nor in priority problems. For instance, in the ultimatum game, the model explains why the first player has to offer a rather large amount of money to the second player, but it does not specify how large this amount should be. Only by playing the game a number of times with the same opponent can a player discover how much he has to offer.

4 Our proposed model: priority awareness

4.1 Fairness boundaries

In order to obtain a fair distribution of rewards over agents with priorities, we introduce a new notion of fairness: priority awareness. We define two boundaries in our model. To keep reward distributions within the boundaries, we introduce a parameter $\alpha \in [0, 1]$, called the *greediness* parameter.[1] With $\alpha = 0$, all agents are satisfied when the total reward is shared evenly (i.e., everybody gets the same reward). With $\alpha = 1$, agents want a reward that is proportional to their priority value. Note that situations where $\alpha < 0$ (someone gets less than equal share even though his priority is not lower) as well as situations where $\alpha > 1$ (someone gets more than a priority-based share) are common in the world around us. However, these situations are generally considered to be unfair. Reward distributions obtained by setting the α parameter to valid values can therefore be best described as being 'not definitely unfair'; i.e., we do not allow reward distributions that can immediately be marked as unfair.

4.2 Formal model

More formally, in our model, we distinguish n different agents i, each with a priority value p_i and a given reward R_i. The priority values are defined such that $\sum_{i=1}^{n} p_i = 1$ and $\forall i : 0 < p_i \leq 1$. We introduce the greediness parameter $\alpha \in [0, 1]$ and define the following class of fairness functions $f_\alpha(i)$, applied to agents i:

$$f_\alpha(i) = \frac{1}{n} + \alpha \left(p_i - \frac{1}{n} \right) \tag{1}$$

Now, we say that agents i and j have a fair share with respect to each other and the parameter α if and only if their rewards R_i and R_j satisfy the following equation:

$$\frac{R_i}{f_\alpha(i)} = \frac{R_j}{f_\alpha(j)} \tag{2}$$

To simulate the flexibiity of human fairness, our model needs to tolerate a (small) range of α-values. Therefore, we need to decide on upper and lower bounds for α, i.e. $\alpha \in$

[1] In the remainder of this text, we assume that every agent uses the same α-value. The model can easily be equipped with 'personal' values for α, similar to the homo egualis model.

$[\alpha_{\min}, \alpha_{\max}]$. These upper and lower bounds then specify how tolerant we are with respect to differences in reward. Using these bounds, we define a *score function* for any pair of agents i and j as follows:

$$s(i,j) = \begin{cases} 1 & \exists \alpha \in [\alpha_{\min}, \alpha_{\max}] : \frac{R_i}{f_\alpha(i)} = \frac{R_j}{f_\alpha(j)} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Thus, two agents i and j yield a $s(i,j)$ of 1 if the reward of agent i is within an interval around the reward of individual j, as specified by the minimum and maximum allowed α-values. The utility for each agent $i \in \{0, 1, \ldots, n\}$ is now determined as:

$$u_i = \frac{1}{n-1} \sum_{j=1, i \neq j}^{n} s(i,j) \tag{4}$$

Thus, each agent scores 1 for every other agent that has a fair reward compared to this agent's reward. In the worst case, an agent obtains a utility of 0. In the best case, an agent has a fair reward compared to all other agents and therefore 'scores' $n-1$ times; we divide the result, $n-1$, by $n-1$, yielding a utility of 1. Thus, $u_i \in [0,1]$ for all i. Now, we define the utility of the whole group of agents as:

$$U = \min_i u_i \tag{5}$$

In order to maximize this group utility, the agents must make sure that every agent is treated in a fair way (according to the model).

5 Initial model validation

To validate our model, we have performed two different experiments with human subjects. Both experiments were performed with 50 subjects, with different subjects for each experiment. Respondents were asked to mark which of the given answers they considered to be fair. We categorized each possible answer as either within or outside our model's boundaries. The first experiment concerned the fruit store test described in Section 3.2. The second experiment contained two questions, both examples of problems in which priorities play a role (see Table 1).

In the first experiment, with a setting of $p = 0.6$, 42 people place the robot somewhat closer to the grapes than to the lychees. Five people place the robot at the grapes; three people place the robot in the middle. Thus, 45 out of 50 positions selected are within our model's boundaries. When the probability that grapes are requested is increased to 0.9, the robot is placed closer to the grapes by everyone who has not placed it at the grapes already. Once again, 45 out of 50 answers given were within our model's boundaries. Thus, in the first experiment, our model predicted 90% of the answers.

In the second experiment, 180 answers were given (respondents could select multiple answers, see Table 2). Of these 180 answers, 164 were within our model's boundaries. Thus, in the second experiment, our model predicted 91% of the human answers.

1. A mail company offers priority stamps and regular stamps. Priority stamps are twice as expensive. How should the delivery time of priority mail compare to that of regular mail? a. First deliver each priority mail and after that, if there is time, deliver regular mail. b. Deliver priority mail twice as fast as regular mail. c. Deliver priority mail faster than regular mail if possible. d. Deliver all mail just as fast. e. Deliver regular mail faster than priority mail.
2. You just had dinner with a large group of people. You all ate and drank different things, so everybody's bill is different. The waiter now brings one bill. How should the bill be distributed? a. The person that has used the most expensive food and drinks pays everything. b. Everybody pays his/her own bill. c. Everybody pays so that the amount paid is somewhat proportional to the expenses done. d. Everybody pays the same. e. People that have eaten less expensive food pay more.

Table 1. Some additional experiments in human fairness.

Question	Answ. a $\alpha > 1$	Answ. b $\alpha = 1$	Answ. c $\alpha \in \langle 0, 1 \rangle$	Answ. d $\alpha = 0$	Answ. e $\alpha < 0$	Sum
1	20%	48%	30%	1%	0%	79
2	0%	39%	34%	28%	0%	101
					Total	180
					Correct	164
					Cover	**91%**

Table 2. The results for the experiments presented above.

6 Conclusion and future work

In this paper, we have argued why fairness is an important issue for designers of many multi-agent systems. We discussed two existing models of human fairness and introduced a third model, called *priority awareness*. In contrast to the two existing models, this model is able to explain and predict human behavior in problems where priorities play an important role.

It is important to realize that this model does not say anything about the optimality of a reward distribution. For instance, a situation in which everyone gets a reward of 0 may very well be fair, but is obviously sub-optimal. In practical situations, people would like to obtain an *optimally fair* reward distribution, meaning that (a) the reward is distributed in a fair way and (b) the total reward is maximized, i.e., the distribution is Pareto-optimal [14]. In our model, this may entail selecting the most optimal α-value. However, more research is needed to assess how fairness and optimality can be balanced.

In the near future, we will conduct more experiments with human subjects in order to support and refine our priority-aware fairness model. Moreover, we will investigate the theoretical properties of solutions our model considers as fair; for instance, we will determine whether a fair policy is able to deal with other policies that try to exploit fair agents. Finally, we will determine how fairness can actually be engineered into a multi-agent system. Our current model already allows us to determine utility values for agents, which can for instance be used to learn a fair policy.

Acknowledgments. The research reported here is partially funded by the Breedtestrategie programme of the Universiteit Maastricht. We thank our colleagues for their constructive criticism and our respondents for their cooperation.

References

1. Erev, I., Roth, A.E.: Predicting how people play games with unique, mixed strategy equilibria. American Economic Review **88**(2/3) (1998) 848–881
2. Russell, S., Norvig, P.: Artificial Intelligence, A Modern Approach. 2nd edn. Prentice Hall (1995)
3. Bowles, S., Boyd, R., Fehr, E., Gintis, H.: Homo reciprocans: A Research Initiative on the Origins, Dimensions, and Policy Implications of Reciprocal Fairness. Advances in Complex Systems **4**(2/3) (1997) 1–30
4. Fehr, E., Schmidt, K.: A Theory of Fairness, Competition and Cooperation. Quarterly Journal of Economics **114** (1999) 817–868
5. Gintis, H.: Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction. Princeton University Press (2001)
6. Shoham, Y., Powers, R., Grenager, T.: If Multi-Agent Learning is the Answer, What is the Question? Journal of Artificial Intelligence. **To appear** (2006)
7. Tuyls, K., Nowe, A.: Evolutionary Game Theory and Multi-Agent Reinforcement Learning. The Knowledge Engineering Review **20** (2005) 63–90
8. Nash, J.: The Bargaining Problem. Econometrica **18** (1950) 155–162
9. Gale, J., Binmore, K.G., Samuelson, L.: Learning to be Imperfect: The Ultimatum Game. Games and Economic Behavior **8** (1995) 56–90
10. Binmore, K.: Natural Justice. Oxford University Press (2005)
11. Rabin, M.: Incorporating Fairness into Game Theory and Economics. American Economic Review **83** (1993) 1281–1302
12. Fehr, E., Fischbacher, U.: The Nature of Human Altruism. NATURE **425** (23 October 2003) 785–791
13. de Jong, S., Tuyls, K., Sprinkhuizen-Kuyper, I.: Robust and Scalable Coordination of Potential-Field Driven Agents. In: Proceedings of IAWTIC/CIMCA 2006, Sydney. (2006)
14. Verbeeck, K.: Coordinated Exploration in Multi-Agent Reinforcement Learning. PhD thesis, Vrije Universiteit Brussel (2004)