Evolution of Collective Fairness in Hybrid Populations of Humans and Agents

Fernando P. Santos,[1,2,4] **Jorge M. Pacheco,**[3,4] **Ana Paiva,**[1] **Francisco C. Santos**[1,4]

[1]INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, IST-Taguspark, 2744-016 Porto Salvo, Portugal
[2]Princeton University, Department of Ecology and Evolutionary Biology, 08544 NJ, USA
[3]CBMA and Departamento de Matemática e Aplicações, Universidade do Minho, 4710-057 Braga, Portugal
[4]ATP-group, P-2744-016 Porto Salvo, Portugal
fernando.pedro@tecnico.ulisboa.pt

Abstract

Fairness plays a fundamental role in decision-making, which is evidenced by the high incidence of human behaviors that result in egalitarian outcomes. This is often shown in the context of dyadic interactions, resorting to the Ultimatum Game. The peculiarities of group interactions – and the corresponding effect in eliciting fair actions – remain, however, astray. Focusing on groups suggests several questions related with the effect of group size, group decision rules and the interrelation of human and agents' behaviors in hybrid groups. To address these topics, here we test a Multiplayer version of the Ultimatum Game (**MUG**): proposals are made to groups of Responders that, collectively, accept or reject them. Firstly, we run an online experiment to evaluate how humans react to different group decision rules. We observe that people become increasingly fair if groups adopt *stricter* decision rules, *i.e.*, if more individuals are required to accept a proposal for it to be accepted by the group. Secondly, we propose a new analytical model to shed light on how such behaviors may have evolved. Thirdly, we adapt our model to include agents with fixed behaviors. We show that including hardcoded Pro-social agents favors the evolutionary stability of fair states, even for *soft* group decision rules. This suggests that judiciously introducing agents with particular behaviors in a population may leverage long-term social benefits.

Introduction

Contrarily to what mainstream economics postulates, individuals often act in favor of others and develop other-regarding preferences, seemingly deviating from stylized models of rationality (Camerer 2003). In fact, individuals reveal a strong bias towards fairness, incurring sizable costs to favor others or punishing those that behave unfairly (Fehr and Fischbacher 2003). Unveiling the origin of those behaviors is a multidisciplinary challenge. Understanding how such actions will be altered when interacting with artificial agents – and knowing how to design multiagent systems accordingly – remains, today, a conundrum in Artificial Intelligence (AI) (Parkes and Wellman 2015).

In this context, the Ultimatum Game (UG) stands as a simple interaction paradigm that is capable of capturing the essential clash between rationality and fairness (Güth, Schmittberger, and Schwarze 1982). In its original form, two players interact in two distinct roles: Proposer and Responder. The Proposer is endowed with some resource and has to propose a division to the second player, the Responder. The Responder has to accept or reject. If the proposal is rejected, none of the players earns anything. If accepted, they will divide the resource as it was proposed. A fair outcome is usually defined as an egalitarian division, in which both the Proposer and the Responder earn a similar reward.

The minimalism of the UG is convenient to attempt predicting, mathematically, how humans will play it. A first approach assumes that each agent is rational and seeks to maximize individual rewards. In this case, thinking in a backward fashion, one may realize that the Responder should always accept any offer; the Proposer should, thereafter, offer the minimum possible amount. Indeed, this line of thought gives an intuition for the *sub-game perfect equilibrium* of the UG: low offers by Proposers and low acceptance thresholds by Responders (Osborne 2004). These predictions are, however, misleading. A vast number of works report experiments in which people behave differently from the rational prediction. Humans tend to reject low proposals and to make high offers (Güth, Schmittberger, and Schwarze 1982; Fehr and Fischbacher 2003; Camerer 2003).

While the UG constitutes a quite general interaction paradigm, there are several human and human-agent interactions not covered by such a pairwise interaction model. Take the case of democratic institutions, economic and climate summits, collective bargaining, markets, auctions, or the ancestral activities of proposing divisions regarding the loot of group hunts and fisheries. All those examples go beyond a pairwise interaction. Indeed, there is a growing interest in analyzing multiplayer versions of the UG (Grimm et al. 2017; Takesue, Ozawa, and Morikawa 2017). A simple extension may turn the UG adequate to study a wide variety of ubiquitous formats of people encounters: in the Multiplayer UG (MUG) proposals are thereby made to groups and the groups should decide, through suffrage, about its acceptance or rejection (Santos et al. 2015).

In AI, a multiplayer version of the UG can further constitute a convenient interaction model to test the emergence of fairness in groups composed by humans and agents (Correia et al. 2018), which can be of natural interest for areas such as multiagent resource allocation (Chevaleyre et al. 2006) or multilateral automated bargaining (Jennings et al. 2001).

Additionally, with the possibility that humans and artificial agents co-exist in populations it is, more than ever, fundamental to understand how human behavior changes when taking part in collective decisions with agents. Studies show that fairness in the UG relies on a complex neural architecture: when facing unfair proposals by other humans, the areas of the brain that get activated are those associated with negative emotional states, such as anger and disgust (Sanfey et al. 2003). Introducing machines and artificial agents in the game may thus result in different responses, as the attribution of causality shifts (Blount 1995). Designing artificial agents that incorporate the mechanisms responsible for the levels of fairness observed in human interactions is non-trivial. Will humans infer causes and assign responsibilities to artificial agents? Will agents blame humans (or other agents) for unfair behaviors?

From a population dynamics standpoint, understanding the long-term evolution of behaviors when playing the MUG can help designing agents that perform better in the long-run, in human-agent interactions – a purely selfish agent that always offers close to nothing to a human Responder will naturally have a lot of proposals rejected (de Jong and Tuyls 2011). Additionally, understanding the evolutionary dynamics of the MUG can help designing agents that are able to steer collective human behavior towards fair, cooperative or, in general, socially desirable states (Paiva, Santos, and Santos 2018). In this realm, we shall highlight a recent result showing how noisy bots aid coordination in populations of humans and agents (Shirado and Christakis 2017).

Here we study the dynamics of Multiplayer Ultimatum Games both experimentally and theoretically. First, we conduct a behavioral experiment involving humans on Amazon Mechanical Turk (AMT) (Mason and Suri 2012). We ask people to play the MUG and we record their strategies, testing different responses when subject to different group decision rules. Secondly, we propose a new theoretical model – inspired in Evolutionary Game Theory (EGT) (Weibull 1997) – to understand the experimental results. The model that we propose will also be extended to incorporate a small set of agents with fixed behaviors (hard-coded agents). With these experiments and model, we seek to answer:

1. How do humans play MUGs and how do human proposals – and ensuing fairness – depend on group decision rules?

2. How do different group decision rules impact the evolutionary dynamics of strategy adoption in MUG?

3. What is the impact, in the evolutionary dynamics of the MUG, of including hardcoded Pro-social agents in a hybrid human-agent population?

As detailed below, we confirm experimentally that particular group decision criteria have an important role in the strategies adopted by people: *stricter* group acceptance rules induce individuals to be fairer. Furthermore, through our theoretical model, we can obtain a similar conclusion: we observe an increase in the basin of attraction towards Pro-social strategies (*i.e.*, proposing and accepting high values) whenever we test *stricter* group decision rules. Finally, and importantly, we conclude that fairness can emerge even for *softer*

decision rules, if a small number of agents with hardcoded Pro-social behavior is introduced in the population. This result highlights the potential of designing intelligent agents able to support and generate pro-social behaviours among Humans, envisioning a broader *pro-social computing* effort within AI (Paiva, Santos, and Santos 2018).

Related work

Several models try to justify the evolution of fairness in the context of Ultimatum Games. Resorting to EGT, Nowak et al. suggested that if Proposers are able to get pieces of information about previous actions of the opponents, then it is worth for the Responders to cultivate a fierce reputation (Nowak, Page, and Sigmund 2000). This way, Proposers would offer more to Responders that are used to reject low offers and this naturally incentivizes Responders to nurture an intransigent reputation by rejecting unfair offers. Other models attribute the evolution of fairness to repeated interactions over time (Van Segbroeck et al. 2012). A slight different approach suggests that fair Proposers and Responders may emerge due to the topological arrangement of their network of contacts: if individuals are arranged in lattices (Page, Nowak, and Sigmund 2000) clusters of fairness may emerge. Rewiring links with neighboring agents was also noted to promote fairness in the UG, specially among agents that resort to learning algorithms to adapt their strategies over time (de Jong, Uyttendaele, and Tuyls 2008). While these results apply to the two-person UG, their generatization to a multiplayer setting is not straightforward.

The formalization of the MUG that we follow here was proposed in (Santos et al. 2015). A more recent study resorts to reinforcement learning (the Roth-Erev algorithm) to show that higher proposals are likely to emerge when stricter rules are considered (Santos et al. 2016). The overall dynamics of strategy adoption that pertains different decision rules, together with the impact of including agents with a predetermined behavior, have, however, remained unaddressed issues, to date. An alternative multiplayer formulation of the UG was proposed in (Takesue, Ozawa, and Morikawa 2017).

On the one hand, the topic of this work is related with the broad agenda of multiagent resource allocation (Chevaleyre et al. 2006), particularly the endeavor of fairly allocating divisible goods (often called cake-cutting). Multiagent resource allocation constitutes a common ground between economics and computer science: while economics typically asks what constitutes a good division (Chevaleyre et al. 2006), computer science tries do advance algorithms to reach divisions that satisfy desirable properties – such as proportionality, envy-freeness (Segal-Halevi, Hassidim, and Aumann 2015), strategyproofness (Brânzei et al. 2016) or truthfulness (Chen et al. 2010). Here, it is noteworthy how the complexity of algorithms increases when moving from the two-agent scenario – where the simple proportional and envy-free "cut-and-choose" algorithm can be applied – to the *n*-person scenario. By moving from the two-person UG to the multiplayer UG we face several challenges alike, particularly related with the many possible ways of mapping individual interests into collective decisions. However, some aspects differentiate our work from traditional research on

cake-cutting algorithms: 1) here we assume that all agents have the same valuation function and the goods to be divided are homogeneous; 2) we focus on the strategic dimension of the division process and the contrast between individual gain and social welfare (equality); 3) lastly and importantly, rather than building an algorithm to divide a good in a desirable way, we try to predict the allocations that are more likely to emerge given the decentralized interaction of adaptive self-regarding agents.

Finally, we shall highlight recent works that use similar methodologies as the one employed here, namely EGT to study behavioral dynamics. The Han et al. proposed an analytical model – with a similar basis as the one we provide below – to show that if the costs of committing to cooperation are smaller than the costs of cooperating itself, then cooperation is able to evolve (Han, Pereira, and Lenaerts 2017). The same authors used an EGT model to show how centralized commitments can be achieved in the context of Public Goods Games, when they are managed by a centralized entity (Han et al. 2017). Recently, EGT was employed to evaluate the emergence of cooperation through norms and reputations (Santos, Pacheco, and Santos 2018) and the stability of normative systems (Morales et al. 2018).

Model and Methods

Multiplayer Ultimatum Game

Let us start by describing the Multiplayer Ultimatum Game (MUG), the interaction paradigm used throughout this paper. We assume that an offer p is made by a Proposer to a group of $N - 1$ Responders. Each of the Responders states acceptance or rejection. Overall acceptance by the group depends on having a minimum number M of individual acceptances. In this case, the Proposer keeps what she did not offer $(1-p)$ and the offer is evenly divided by all the Responders $(p/(N - 1))$; otherwise, if the number of acceptances remains below M, the proposal is rejected by the group and no one earns anything.

The payoff function of a Proposer i, with strategy p_i, facing a group of Responders with strategies $q_{-i} = \{q_j : j = 1, ..., N-1 \wedge j \neq i\}$ reads as $\Pi_P(p_i, q_{-i}) = (1-p_i)a_{p_i,q_{-i}}$, where $a_{p_i,q_{-i}}$ summarizes the group acceptance of the proposal made by agent i, p_i, standing as

$$a_{p_i,q_{-i}} = \begin{cases} 1, & \text{if } \sum_{q_j \in q_{-i}} \Theta(p_i - q_j) \geq M. \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$\Theta(x)$ is the Heaviside unit step function, having value 1 whenever $x \geq 0$ and 0 otherwise. This way, $\Theta(p_i - q_j) = 1$ if agent j accepts agent i proposal.
Similarly, the payoff function describing the gains of a Responder belonging to a group with a strategy profile $q_{-j} = \{q_k : k = 1, ..., N - 1 \wedge k \neq j\}$ listening to a Proposer j with strategy p_j, is given by $\Pi_R(p_j, q_{-j}) = \frac{p_j}{N-1}a_{p_j,q_{-j}}$.

The interesting values of M range between 1 and $N - 1$. If $M < 1$ all proposals would be accepted (resulting in the so-called Dictator game) and having $M > N - 1$ would mean unconditional rejection. We shall add that the *subgame perfect equilibrium* of this game is, as in the tradition

two-person UG, composed by the lowest possible values of p and q (Santos et al. 2016).

Behavioral experiments

In order to test the most likely values of p and q used by people, we recruited individuals using Amazon Mechanical Turk (AMT) (Mason and Suri 2012) – for examples of AMT usage in the community, see (Azaria, Aumann, and Kraus 2012; Rosenfeld and Kraus 2015). For each combination of $M = \{1, N - 1\}$ with $N = \{5, 10\}$ we collected 100 pairs of strategies (p, q). We followed the strategy method (de Melo, Marsella, and Gratch 2018), informing people about the game rules, group size and group decision rule, and asking participants to submit their proposal as a Proposer (p) and the minimum acceptable offer as a Responder (q). We provide an eliminatory test, used to guarantee that individuals actually understood the rules of the MUG. Only after passing this test people were able to introduce the real strategies to play the MUG. The average values of the strategies presented in the tests were similar in the 4 treatments. Each participant was paid a base amount of 0.4 USD and a bonus value between 0 and 0.6 USD was endowed depending on the performance when playing the MUG. After all strategies were collected for a given treatment, random groups were assembled and the strategies introduced by participants were used to calculate the resulting bonus. Each participant was paid accordingly and strategies were recorded.

Analytical model

In the analytical Evolutionary Game Theoretical (EGT) model, we consider a simplified version of the Multiplayer Ultimatum Game (Mini-MUG), inspired by a similar simplification of the UG proposed in (Nowak, Page, and Sigmund 2000). We constrain the values of proposal (p) and acceptance threshold (q) to two values, named high (h) and low $(l, 0 < l < h < 1)$. This way, defining a strategy as a duple (p, q) of proposal p and acceptance threshold q, individuals select one out of four possible options: *Pro-social* (h, h), *Asocial* (l, l), *Generous* (h, l) and *Paradoxical* (l, h). Following previous works (Nowak, Page, and Sigmund 2000), we focus on the evolutionary dynamics between three of those strategies: Pro-social (P), Asocial (A) and Generous (G).

Once a group with size N is assembled, one individual is randomly selected to be the Proposer. Each individual has a probability $1/N$ of being the Proposer and $(N - 1)/N$ to play as a Responder. This said, the average payoff of an individual playing strategy P in a group where i' other individuals adopt strategy P, j' individuals adopt A and $N-i'-j'-1'$ individuals adopt G, is

$$\Pi_P = \frac{1-h}{N} + \frac{j'}{N}\frac{A(i'+1)l}{N-1} + \frac{N-j'-1}{N}\frac{h}{N-1}, \quad (2)$$

where, as before, $A(i') = \Theta(N - 1 - i' - M)$ is the Heaviside unit step function: $\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \geq 0$. This means that $A(i')$ yields 1 if a proposal l is accepted (*i.e.*, the number of acceptances is at minimum M). The right-hand side of Π_P is constituted by three terms: 1) the amount a Pro-social individual earns as a Proposer: $1-h$,

a quantity that is always accepted; 2) what the individual earns as the Responder of a low offer that may (or may not) be accepted and 3) what that individual earns as a Responder to a high offer. Similarly, when facing the same group composed by i' Ps and j' As, the average payoff of an individual using A is

$$\Pi_A = (\frac{1-l}{N} + \frac{j'}{N}\frac{l}{N-1})A(i') + \frac{N-j'-1}{N}\frac{h}{N-1}, \quad (3)$$

and the average payoff of an individual G stands as

$$\Pi_G = \frac{1-h}{N} + \frac{j'}{N}\frac{A(i')l}{N-1} + \frac{N-j'-1}{N}\frac{h}{N-1}. \quad (4)$$

Evolutionary Dynamics: Within a finite population of size Z, groups of size N are randomly sampled and individuals will accumulate payoff according with the payoff functions defined above. In a population composed by i individuals with strategy P, j individuals with strategy A and $Z-i-j$ individuals with strategy G, the probability of sampling a group with i', j' and $N-i'-j'$ individuals using, respectively, strategies P, A and G, is given by the multivariate hypergeometric distribution $S(i,j,i',j',N,Z) = \binom{i}{i'}\binom{j}{j'}\binom{Z-i-j}{N-i'-j'}/\binom{Z}{N}$. This said, the fitness (or average payoff) of an individual using strategy P in a population state (i,j) is given by

$$f_P = \sum_{i'=0}^{N-1}\sum_{j'=0}^{N-1-i'}\Pi_P(i',j')S(i-1,j,i',j',N-1,Z-1). \quad (5)$$

The fitness of individuals using strategy A and G are similarly constructed and are not explicitly written for the sake of space. These fitness functions allow defining a stochastic process over all the possible population states, defined by each possible combination of strategies, (i,j). For that, we assume that individuals will change their strategy through social learning, imitating the strategies of the best performing (*i.e.*, with higher fitness) individuals (Nowak, Page, and Sigmund 2000; Han, Pereira, and Lenaerts 2017; Han et al. 2017). The transition probability from a state (i,j) to a state $(i+1,j)$ is calculated as the probability that an A or G individual imitates a P individual. As an example, the probability that A imitates P is given by the probability of selecting an A ($\frac{j}{Z}$), times the probability of selecting a P ($\frac{i}{Z-1}$), times the probability that A imitates P ($p_{A\to P}$). All together, the probability of having one more P is,

$$T_P^+ = \frac{(1-\mu)i}{Z-1}(\frac{j}{Z}p_{A\to P} + \frac{Z-j-i}{Z}p_{G\to P}) + \frac{\mu(Z-i)}{2Z}. \quad (6)$$

$p_{X\to Y}$, the probability that an individual with strategy X imitates one with strategy Y, is calculated resorting to the sigmoid function $p_{X\to Y} = (1 + e^{\beta(f_X - f_Y)})^{-1}$, which allows defining this imitation step as a probabilistic event event, prone to errors (Traulsen, Nowak, and Pacheco 2006). The value β controls the selection strength, *i.e.*, the extent to which the evolutionary process depends on the fitness difference between individuals: high β implies that small deviations in fitness differences strongly impact the probability of imitation; low β means that imitation depends weakly on fitness differences, which in social systems may be due to

imitation errors or the incorrect assessment of peers' performance. We set $\beta = 1$ throughout this work, yet similar results were verified for $\beta = \{0.5, 2\}$. Also, an exploration term is introduced: we consider that imitation occurs with probability $1-\mu$ and, with probability μ, an individual spontaneously adopts other strategy, without resorting to any fitness-driven heuristic: with probability $\frac{\mu j}{Z}\frac{1}{2}$ an individual with strategy A mutates to strategy P; with probability $\frac{\mu(Z-i-j)}{Z}\frac{1}{2}$ an individual with strategy G mutates to strategy P; when summed, we obtain the term $\frac{\mu(Z-i)}{2Z}$ in Eq. (6). The transition probability from (i,j) to $(i-1,j)$ is given by

$$T_P^- = \frac{(1-\mu)i}{Z}(\frac{j}{Z-1}p_{P\to A} + \frac{Z-j-i}{Z-1}p_{P\to G}) + \frac{\mu i}{Z}. \quad (7)$$

Following an analogous reasoning for the update of strategy A, we can define the transition probability from a state (i,j) to a state $(i,j+1)$ as

$$T_A^+ = \frac{(1-\mu)j}{Z-1}(\frac{i}{Z}p_{P\to A} + \frac{Z-j-i}{Z}p_{G\to A}) + \frac{\mu(Z-j)}{2Z}, \quad (8)$$

and the transition probability from a state (i,j) to a state $(i,j-1)$ as

$$T_A^- = \frac{(1-\mu)j}{Z}(\frac{i}{Z-1}p_{A\to P} + \frac{Z-j-i}{Z-1}p_{A\to G}) + \frac{\mu j}{Z}. \quad (9)$$

Gradient of Selection and Stationary Distribution: The probabilities defined above allow us to define a Markov Chain whose transition matrix T is defined as

$$\begin{aligned}
T_{(i,j)\to(i+1,j)} &= T_P^+(i,j) \\
T_{(i,j)\to(i-1,j)} &= T_P^-(i,j) \\
T_{(i,j)\to(i,j+1)} &= T_A^+(i,j) \\
T_{(i,j)\to(i,j-1)} &= T_A^-(i,j) \\
T_{(i,j)\to(i,j)} &= 1 - T_P^+(i,j) - T_P^- - T_A^+ - T_A^- \\
T_{X\to Y} &= 0, \text{otherwise.}
\end{aligned} \quad (10)$$

We can now determine the average fraction of time that the population spends in each state, after a long period of time elapses. For that, we study the limiting distribution of the Markov process whose transition probabilities are given by T. If matrix T defines a finite-state discrete-time Markov Chain with a single recurrent class (which is the case when $\mu > 0$), then we can calculate the average fraction of time spent on each state (after a long period of time and independently of the starting state) by solving the equation

$$\pi = \pi T. \quad (11)$$

π is the stationary distribution of this process and we can calculate it by noting that π is the eigenvector of T associated with its highest ($= 1$) eigenvalue. We use π to calculate the average usage of each strategy in Fig. 2 and 4. After arranging all the possible states (i,j) in a 2-dimensional simplex (see Fig. 3 and 5) we can visually grasp the stationary distribution by coloring each state with the corresponding value in π (normalized). We assume that the darker the state
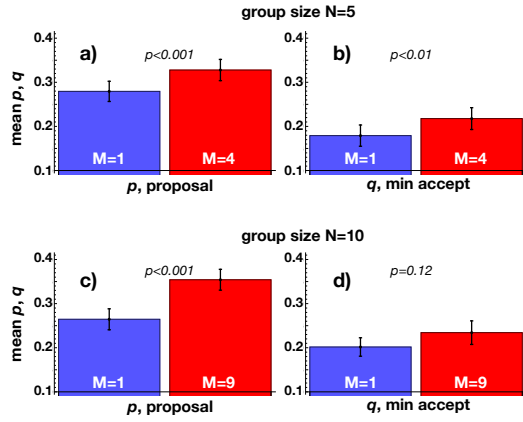
Figure 1: Effect of M on the average proposal and acceptance threshold. Error bars represent the standard error of the mean. A Mann-Whitney U test was conducted to compare the average offer (p) and acceptance threshold (q) in $M = 1$ and $M = N - 1$ conditions, when the group size is $N = 5$ or $N = 10$. **a)** Considering $N = 5$, there was a significant difference in the average offer (p) for $M = 1$ (Mean=0.279, SD=0.117) and $M = 4$ (Mean=0.328, SD=0.123); U=3846, n1=n2=100, $p < 0.001$. **c)** Considering $N = 10$, there was a significant difference in the average offer (p) for $M = 1$ (Mean=0.265, SD=0.122) and $M = 9$ (Mean=0.354, SD=0.121) conditions; U=2724.5, n1=n2=100, $p < 0.001$. These results suggest that people care about group acceptance rule when making an offer: increasing M induces higher – and fairer – proposals. Regarding the average acceptance threshold (q), for $N = 5$ **b)** there was a significant difference for $M = 1$ (Mean=0.179, SD=0.123) and $M = 4$ (Mean=0.218, SD=0.127); U=3913.5, n1=n2=100, $p<0.01$. For $N=10$ **d)** there was not a significant difference in q for $M=1$ (Mean=0.202, SD=0.107) and $M=9$ (Mean=0.234, SD=0.137) conditions; U=4373.5, n1=n2=100, $p=0.124$. These results suggest that M can also play a role (yet less significantly) in the acceptance threshold of individuals.

is the more time is spent there. Also, we can visualize the most likely evolutionary path, starting in any state (i, j), as a vector field defined by the so-called gradient of selection,

$$\vec{g}(i, j) = [T_P^+(i, j) - T_P^-(i, j), T_A^+(i, j) - T_A^-(i, j)]. \quad (12)$$

Hardcoded agents: We extend the base model, defined before, to account for the existence of a fixed number k of Pro-social hardcoded agents that maintain their behavior over time, never changing their strategy through social learning or exploration. This implies changing the transition probabilities T_P^- and T_A^+ (named T_{kP}^- and T_{kA}^+) which now become,

$$T_{kP}^- = \frac{(i - k)\Theta(i - k)}{Z}\left((1 - \mu)\left(\frac{j}{Z - 1}p_{A \to P}\right.\right.$$
$$\left.\left. + \frac{Z - j - i}{Z - 1}p_{G \to P}\right) + \mu\right), \quad (13)$$
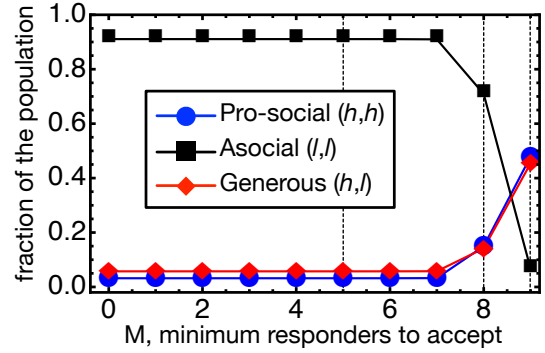

Figure 2: Results from the evolutionary game theoretical model: average values of proposal and acceptance threshold increase with M. Whereas almost all the individuals adopt the Asocial strategy (black squares) for low values of M, we note a significant increase in the usage of Pro-social (blue circles) and Generous (red diamonds) strategy for $M = 8$ and $M = 9$. As in the experiments (Fig. 1) stricter group decision rules induce fairer agents. The vertical dashed lines represent the scenarios detailed in Fig. 3. Parameters used: $h = 0.6$, $l = 0.1$, $Z = 100$, $\mu = 0.02$, $N = 10$.

$$T_{kA}^+ = \frac{(i - k)\Theta(i - k)}{Z}\left(\frac{j(1 - \mu)}{Z - 1}p_{P \to A} + \frac{\mu}{2}\right)$$
$$+ \frac{Z - j - i}{Z - 1}\left(\frac{j(1 - \mu)}{Z - 1}p_{G \to A} + \frac{\mu}{2}\right), \quad (14)$$

As we will see, introducing Pro-social hardcoded agents alters the evolutionary dynamics of fairness in the population, implying long-term social benefit in terms of equality.

Results

Here we present the results obtained from our behavioral experiments employing humans (Fig. 1), from the proposed evolutionary game theoretical model (Fig. 2 and 3) and the theoretical model adapted to consider the existence of hardcoded Pro-social agents (Fig. 4 and 5). We collect a total of 400 strategies, submitted by 400 individuals using Amazon Mechanical Turk (AMT), in 4 different treatments: i) $N = 5$, $M = 1$; ii) $N = 5$, $M = 4$; iii) $N = 10$, $M = 1$; iv) $N = 10$, $M = 9$. In Fig. 1 we present the results for the average proposal (p, left) and acceptance threshold (q, right), considering $N = 5$ (top) and $N = 10$ (bottom), $M = 1$ (blue) and $M = N - 1$ (red). We shall first note that the average proposal, p, significantly increases when M goes from 1 to $N - 1$. For $M = 1$, a single accepting Responder is enough to confer overall group acceptance to a proposal. In this case, we need consensus of rejecting Responders in order for a proposal to be rejected. Contrarily, $M = N - 1$ stands as a stricter consensus threshold, meaning that all the Responders need to accept a proposal for it to be accepted. The average value of proposal increases regardless of the group size ($N = 5$ or $N = 10$). Notwithstanding, the effect is more pronounced for $N = 10$ (average p goes from 0.265 when $M = 1$ to 0.354 when $M = 4$, Fig. 1c) than for $N = 5$ (average p goes from 0.279 when $M = 1$ to 0.328
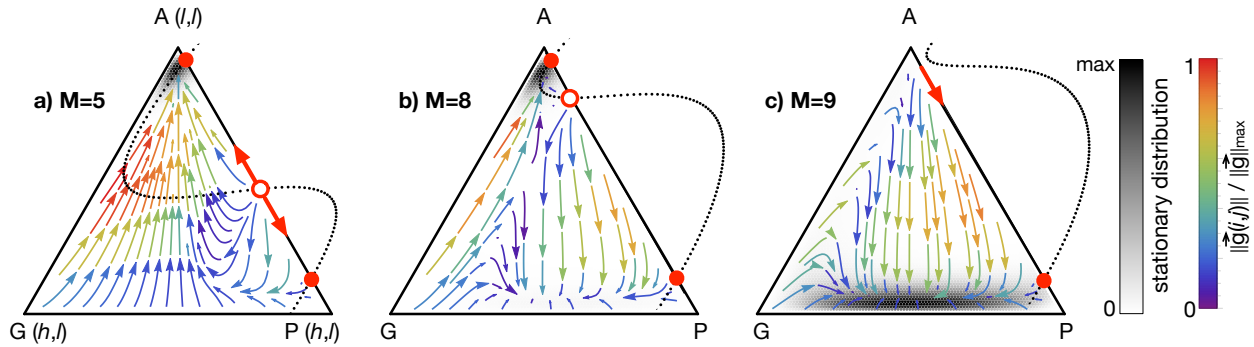
Figure 3: Evolutionary dynamics of Pro-social (P), Asocial (A) and Generous (G) strategies when increasing M. Streamlines represent the gradient of selection, *i.e.*, a picture of the most likely evolutionary paths starting from each possible state, calculated from the vector field generated through Eq. (12); to compare the relative norm of each vector – which translates how high/low transition probabilities are – we color the streamline such that warmer (colder) colors represent higher (lower) magnitudes. The gray scale represents the stationary distribution depicting the average time that the system spends in each possible state (see **Model and Methods** for details). The darker the color, the larger the fraction of time spent. On the right edge of each simplex, we represent the one-dimensional gradient of selection when strategy G is absent from the population (assuming an imaginary Y-axis perpendicular to that edge of the simplex). This allows us to better grasp the dynamics between strategy P and A: whenever the curve assumes positive values, strategy P is favored. The fixed points are represented through full circle (stable fixed point) and empty circles (unstable). We can observe that increasing M from $M = 5$ (panel **a**) to $M = 8$ (**b**) and $M = 9$ (**c**) moves the unstable fixed point between P and A towards vertex A, thus increasing the basin of attraction to the states where P and G prevail. This leads to the results portrayed in Fig. 2, in which increasing M leads to an increase in the fraction of P and G (and a corresponding increase of p and q values). Same parameters as in Fig. 2.
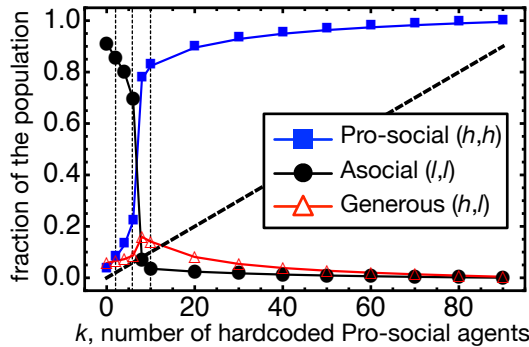

Figure 4: Stationary frequency of strategies P (blue squares), A (black circles) and G (red triangles) after introducing a number k of hardcoded Pro-social agents (*i.e.*, agents keeping the same strategy and not adapting over time). The thick black line represents the number of hardcoded Pro-social agents introduced in the population. There is a significant increase in the number of Pro-social agents, in the long-run – much higher than the number of hardcoded agents directly introduced. The vertical dashed lines represent the scenarios detailed in Fig. 5. Parameters used: $h = 0.6$, $l = 0.1$, $Z = 100$, $\mu = 0.02$, $N = 10$, $M = N/2$.

when $M = 4$, Fig. 1a). Second, we observe that also the average individual Responder threshold, q, increases with M, yet less significantly than what occurs with p (we only find statistical significance in the case $N = 5$, Fig. 1b). In any case, the average values of q always remain below p, as tradi-

tionally occurs in the bilateral UG (Camerer 2003). Finally, we stress that neither p nor q vary monotonously with N. In fact, depending on M, the mean value of p can decrease ($M = 1$) or increase ($M = N - 1$) with N. In a word, fairness increases following the particular decision rule agreed by the group (M) and not with the group size (N).

In Fig. 2 and 3 we show the results from the analytical model detailed in the section **Model and Methods**. Differently from what would be expected theoretically by assuming purely rational agents, in Fig. 2 we observe that increasing M has a positive effect on the average values of p and q, as agents have a higher probability of adopting the Pro-social strategy (high offer and high acceptance threshold) and the Generous strategy (high offer and low threshold).

These results can be understood by assessing the full evolutionary portrait of strategies P, G and A, and how that depends on M. In Fig. 3 we observe that increasing M enlarges the basin of attraction towards states where strategies P and G prevail, thus increasing the fraction of time that agents spend, on average, adopting such strategies.

Finally, we discuss the extension of the model to contemplate the existence of a given number of individuals that adopt a predetermined and fixed behavior, thus providing the opportunity to study, theoretically, the long-term effect of engineered behaviors on a population of adaptive agents. From Fig. 4 we can conclude that introducing a small number of hardcoded Pro-social agents (labeled P') leverages the overall adoption of strategy P to an extent much higher than the fraction of hardcoded agents considered. For instance, including 10% of hardcoded P agents is enough to
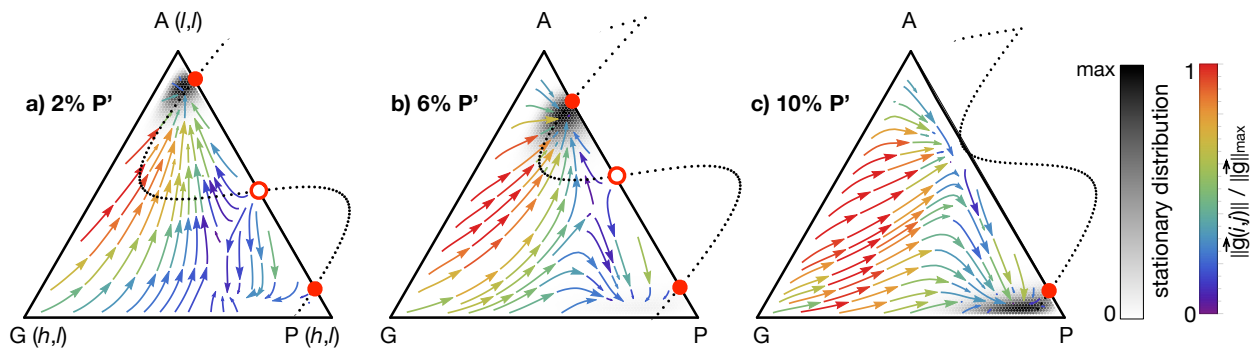
Figure 5: By introducing hardcoded Pro-social agents (P'), we can observe that the closest fixed point to vertex A is moved down and the unstable fixed point is moved up (panels **a** and **b**), eventually merging (panel **c**). Altogether, the basin of attraction to A vanishes, and all states end up leading to configurations in which the Pro-social strategy prevails. Same notation as in Fig. 3 caption. Parameters used: $h = 0.6$, $l = 0.1$, $Z = 100$, $\mu = 0.02$, $N = 10$, $M = N/2$.

obtain, in the long-run, more than 80% of Pro-social individuals (Fig. 4, third vertical dashed line).

This result can, again, be clarified by looking to the evolutionary portrait of the P, A and G strategies (Fig. 5, for the same scenario of Fig. 3a). By introducing hardcoded Pro-social agents (P'), we can observe that the closest fixed point to vertex A is moved right (*i.e.*, towards vertex P) and the unstable fixed point is moved left (*i.e.*, towards vertex A), eventually merging. This results in a bigger basin of attraction towards strategy P and the consequent increase in the long-term adoption of this strategy – much higher than the initial fraction of hardcoded P introduced. This occurs for two main reasons: 1) a larger pool of Pro-social agents is available to imitate and learn with; 2) the ubiquity of Pro-social agents dictates that Asocial agents are less likely to have their unfair proposals accepted.

Discussion and Conclusions

In this work we study a Multiplayer version of the Ultimatum Game (MUG). We perform behavioral experiments with this game in Amazon Mechanical Turk (AMT). Departing from the traditional two-person UG, in the MUG we consider groups of N individuals in which 1 Proposer makes an offer to the remaining $N - 1$ Responders. Two central parameters are added: group size (N) and group acceptance threshold (M). Similarly to the UG, the Responders should always accept any offer ($q = 0$) and the Proposers should offer the minimum possible ($p = 0$). This would naturally occur independently of N and M. We find that, however, increasing M leads to higher average proposals (p) and individual acceptance thresholds (q) (Fig. 1). Stricter decision rules lead to fairer outcomes.

This result can be understood theoretically by means of an evolutionary game theoretical (EGT) model. We show that increasing M significantly augments the population states that most likely lead to configurations where agents adopt Pro-social strategies (offering high proposals). In other words, there is an increase of the basin of attraction towards the Pro-social strategies.

Equipped with a theoretical model that leads to predictions that match the results of behavioral experiments, we then show that introducing a small fraction of hardcoded Pro-social agents in the population results in a long-term increase in the number of Pro-social agents – substantially higher than the number of hardcoded agents introduced. This study suggests, therefore, that judiciously introducing agents with particular behaviors (Pro-social) can have a long-term social benefit by shaping the behaviors of humans and agents (Paiva, Santos, and Santos 2018). We highlight that the model we present and apply in the context of MUG can be tuned to assess the role of hardcoded pro-social agents in the context of other multiplayer games, such as N-Person Snow-Drift (Chiong and Kirley 2012) or Trust Games (Chica et al. 2017) – to name a few.

To conclude, several extensions can be considered in future works. It would be interesting to study the effect of targeted proposals and offers that discriminate between different Responders. Also, while here we consider that group decision rules are exogenously imposed, it would be relevant to test the adherence to a particular group decision rule as part of an agent's strategy. Finally, the reputational effect of not conforming with the majority, by, *e.g.*, being the only one accepting/rejecting a proposal, could be taken into account, given the studies that highlight the important relation between peer pressure, group incentives and group monitoring. Testing this game in behavioral experiments, including robots or virtual agents with predetermined behaviors and studying how humans play along with those, is certainly an additional relevant research avenue to complement the theoretical conclusions that we present here.

References

Azaria, A.; Aumann, Y.; and Kraus, S. 2012. Automated strategies for determining rewards for human work. In *AAAI-12*, 1514–1521. AAAI Press.

Blount, S. 1995. When social outcomes aren't fair: The effect of causal attributions on preferences. *Organ. Behav. Hum. Decis. Process* 63(2):131–144.

Brânzei, S.; Caragiannis, I.; Kurokawa, D.; and Procaccia, A. D. 2016. An algorithmic framework for strategic fair division. In *AAAI-16*, 418–424. AAAI Press.

Camerer, C. 2003. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Chen, Y.; Lai, J. K.; Parkes, D. C.; and Procaccia, A. D. 2010. Truth, justice, and cake cutting. In *AAAI-10*, 756–761. AAAI Press.

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

Chica, M.; Chiong, R.; Kirley, M.; and Ishibuchi, H. 2017. A networked n-player trust game and its evolutionary dynamics. *IEEE Trans. Evol. Comput.*

Chiong, R., and Kirley, M. 2012. Effects of iterated interactions in multiplayer spatial evolutionary games. *IEEE Trans. Evol. Comput.* 16(4):537–555.

Correia, F.; Mascarenhas, S.; Prada, R.; Melo, F. S.; and Paiva, A. 2018. Group-based emotions in teams of humans and robots. In *HRI-18*, 261–269. ACM.

de Jong, S., and Tuyls, K. 2011. Human-inspired computational fairness. *Auton. Agents. Multi. Agent. Syst.* 22(1):103.

de Jong, S.; Uyttendaele, S.; and Tuyls, K. 2008. Learning to reach agreement in a continuous ultimatum game. *J. Artif. Intell. Res.* 33:551–574.

de Melo, C. M.; Marsella, S.; and Gratch, J. 2018. Social decisions and fairness change when people's interests are represented by autonomous agents. *Auton. Agents. Multi. Agent. Syst.* 32(1):163–187.

Fehr, E., and Fischbacher, U. 2003. The nature of human altruism. *Nature* 425(6960):785.

Grimm, V.; Feicht, R.; Rau, H.; and Stephan, G. 2017. On the impact of quotas and decision rules in ultimatum collective bargaining. *Eur. Econ. Rev.* 100:175–192.

Güth, W.; Schmittberger, R.; and Schwarze, B. 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3(4):367–388.

Han, T.; Pereira, L. M.; Martinez-Vaquero, L. A.; and Lenaerts, T. 2017. Centralized vs. personalized commitments and their influence on cooperation in group interactions. In *AAAI-17*, 2999–3005. AAAI Press.

Han, T. A.; Pereira, L. M.; and Lenaerts, T. 2017. Evolution of commitment and level of participation in public goods games. *Auton. Agents. Multi. Agent. Syst.* 31(3):561–583.

Jennings, N. R.; Faratin, P.; Lomuscio, A. R.; Parsons, S.; Wooldridge, M. J.; and Sierra, C. 2001. Automated nego-tiation: prospects, methods and challenges. *Group. Decis. Negot.* 10(2):199–215.

Mason, W., and Suri, S. 2012. Conducting behavioral research on amazon's mechanical turk. *Behav. Res. Methods.* 44(1):1–23.

Morales, J.; Wooldridge, M.; Rodríguez-Aguilar, J. A.; and López-Sánchez, M. 2018. Off-line synthesis of evolutionarily stable normative systems. *Auton. Agents. Multi. Agent. Syst.* 1–37.

Nowak, M. A.; Page, K. M.; and Sigmund, K. 2000. Fairness versus reason in the ultimatum game. *Science* 289(5485):1773–1775.

Osborne, M. J. 2004. *An introduction to game theory*. Oxford University Press New York.

Page, K. M.; Nowak, M. A.; and Sigmund, K. 2000. The spatial ultimatum game. *Proc. R. Soc. Lond. B Biol. Sci.* 267(1458):2177–2182.

Paiva, A.; Santos, F. P.; and Santos, F. C. 2018. Engineering pro-sociality with autonomous agents. In *AAAI-18*, 7994–7999. AAAI Press.

Parkes, D. C., and Wellman, M. P. 2015. Economic reasoning and artificial intelligence. *Science* 349(6245):267–272.

Rosenfeld, A., and Kraus, S. 2015. Providing arguments in discussions based on the prediction of human argumentative behavior. In *AAAI-15*, 1320–1327. AAAI Press.

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(5626):1755–1758.

Santos, F. P.; Santos, F. C.; Paiva, A.; and Pacheco, J. M. 2015. Evolutionary dynamics of group fairness. *J. Theor. Biol.* 378:96–102.

Santos, F. P.; Santos, F. C.; Melo, F. S.; Paiva, A.; and Pacheco, J. M. 2016. Dynamics of fairness in groups of autonomous learning agents. In *AAMAS 2016, Workshops Best Papers*, 107–126. Springer, Cham.

Santos, F. P.; Pacheco, J. M.; and Santos, F. C. 2018. Social norms of cooperation with costly reputation building. In *AAAI-18*, 4727–4734. AAAI Press.

Segal-Halevi, E.; Hassidim, A.; and Aumann, Y. 2015. Envy-free cake-cutting in two dimensions. In *AAAI-15*, 1021–1028. AAAI Press.

Shirado, H., and Christakis, N. A. 2017. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature* 545(7654):370.

Takesue, H.; Ozawa, A.; and Morikawa, S. 2017. Evolution of favoritism and group fairness in a co-evolving three-person ultimatum game. *EPL* 118(4):48002.

Traulsen, A.; Nowak, M. A.; and Pacheco, J. M. 2006. Stochastic dynamics of invasion and fixation. *Phys. Rev. E* 74(1):011909.

Van Segbroeck, S.; Pacheco, J. M.; Lenaerts, T.; and Santos, F. C. 2012. Emergence of fairness in repeated group interactions. *Phys. Rev. Lett.* 108(15):158104.

Weibull, J. W. 1997. *Evolutionary game theory*. MIT press.