Learning to be Fair in Multiplayer Ultimatum Games

(Extended Abstract)

Fernando P. Santos, Francisco C. Santos, Francisco Melo, Ana Paiva
INESC-ID and Instituto Superior Técnico, Universidade de Lisboa
Taguspark, Av. Prof. Cavaco Silva
2780-990 Porto Salvo, Portugal

Jorge M. Pacheco
CBMA and Departamento de Matemática e Aplicações, Universidade do Minho
Campus de Gualtar
4710-057 Braga, Portugal

fernando.pedro@tecnico.ulisboa.pt

ABSTRACT

We study a multiplayer extension of the well-known Ultimatum Game (UG) through the lens of a reinforcement learning algorithm. Multiplayer Ultimatum Game (MUG) allows us to study fair behaviors beyond the traditional pairwise interaction models. Here, a proposal is made to a quorum of Responders, and the overall acceptance depends on reaching a threshold of individual acceptances. We show that learning agents coordinate their behavior into different strategies, depending on factors such as the group acceptance threshold and the group size. Overall, our simulations show that stringent group criteria trigger fairer proposals and the effect of group size on fairness depends on the same group acceptance criteria.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*multiagent systems*; J.4 [**Social and Behavioral Sciences**]: Economics

Keywords

Fairness, Groups, Ultimatum Game, Learning, Multiagent systems

1. INTRODUCTION

The evidence of fairness in decision-making has for long captured the attention of academics and the subject comprises a fertile ground of multidisciplinary research [1]. In this context, UG stands as a simple interaction paradigm that is capable to synthesize the clash between rationality and fairness [2]. In its original form, two players interact acquiring two distinct roles: Proposer and Responder. The Proposer is endowed with some resource and has to propose a division (p) to the second player. After that, the Responder has to state her acceptance or rejection, based on a threshold of acceptance (q). If the proposal is rejected ($p < q$), none of the players earns anything. If the proposal is accepted ($p \geq q$), they will divide the resource as it was proposed: the Proposer keeps $1-p$ and the Responder earns p. A fair outcome is defined as an egalitarian division, $p = 0.5$, in which both the Proposer and the Responder earn a similar reward.

A large set of works address the dynamics of agents playing UG in its original two-player formulation. There are, however, many group interactions in which fairness also plays a fundamental role [7, 4, 3], and which cannot be captured by means of the traditional 2-player UG. Consider, for instance, the cases of democratic institutions, economic and climate summits, collective bargaining, group buying, auctions, or the ancestral activities of proposing divisions regarding the loot of group hunts and fisheries. Here we analyze a group version of UG that allows us address the interactions in which proposals are made to groups and the groups should decide, through suffrage, about its acceptance or rejection [7]. In MUG, proposals are made (p), yet now each of the Responders states acceptance ($p \geq q$) of rejection ($p < q$) and the overall acceptance depends on an aggregate of these individual decisions: if the number of acceptances equals or exceeds a threshold M, the proposal is accepted by the group. In this case, the Proposer keeps what she did not offered, $1-p$, and the offer is evenly divided by all the Responders, $p/(N-1)$; otherwise, if the number of acceptances remain bellow M, the proposal is rejected by the group an no one earns anything. The interesting values of M range between 1 and $N-1$. If $M < 1$ all proposals would be accepted and having $M > N-1$ would dictate unrestricted rejections. If $N = 2$ and $M = 1$ we recover the traditional 2-person UG, described above [2].

2. LEARNING MODEL AND RESULTS

We use the Roth-Erev algorithm [6] to analyse the outcome of a population of learning agents playing MUG in groups of size N. In this algorithm, at each time-step t, the decision-making of each agent k is implemented by a propensity vector $Q_k(t)$ that, as we will see, translates the prob-

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

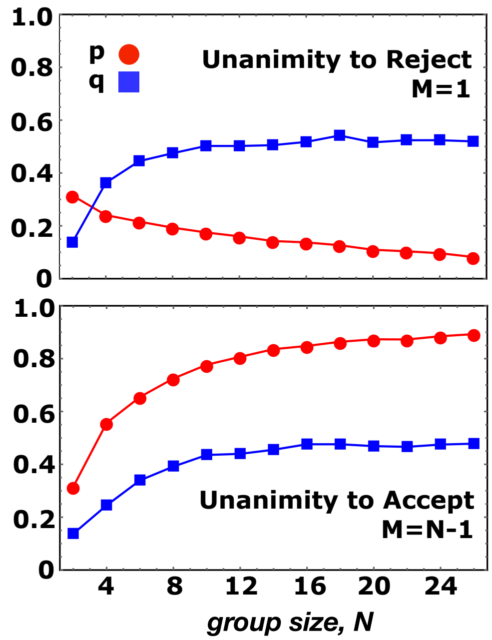

Figure 1: Average values of p and q for different combinations of group sizes, N, and group decision criteria, M when population size $Z = 50$

ability of using each strategy. This vector will be updated considering the payoff gathered in each play. This way, successfully employed actions will have high probability of being repeated in the future. We consider that games take place within a population of size $Z > N$ of adaptive agents. To calculate the payoff of each agent, we sample random groups without any kind of preferential arrangement (well-mixed assumption). We consider MUG with discretised strategies. We round the possible values of p (proposed offers) and q (individual threshold of acceptance) to the closest multiple of $1/D$, where D measures the granularity of the strategy space considered. We map each pair of decimal values p and q into an integer representation, thereafter $i_{p,q}$ is the integer representation of strategy (p, q) and p_i (or q_i) designates the p (q) value corresponding to the strategy with integer representation i. The core of the learning algorithm takes place in the update of the propensity vector of each agent, $Q(t+1)$, after a play at time-step t. Denoting the set of possible actions by $A, a_i \in A : a_i = \{p_i, q_i\}$ and the population size by Z, the propensity matrix, $Q(t) \in \mathbb{R}_+^{Z \times |A|}$ is updated following the base rule

$$Q_{ki}(t+1) = \begin{cases} Q_{ki}(t) + \Pi(p_i, q_i, p_{-i}, q_{-i}) & \text{if } k \text{ played } i \\ Q_{ki}(t) & \text{otherwise} \end{cases} \quad (1)$$

When an agent is called to pick an action, she will do so following the probability distribution dictated by her normalised propensity vector.

Following the traditional game theoretical equilibrium assumptions, the strategy picked by each agent would always be that of unconditional acceptance and minimal offer (subgame perfect equilibrium [5]). Given this prediction, even considering MUG, proposals would always be accepted ($q \rightarrow 0$) and Proposers would always keep the largest share of payoff ($p \rightarrow 0$). This conclusion rules out any possible effect played by group sizes (N) or group acceptance criteria (M). Notwithstanding, through the implemented learning model, we show that larger groups induce individuals to rise their average acceptance threshold (Figure 1). Indeed, it is reasonable to assume that, as the group of Responders grow and as they have to divide the offers between more individuals, the pressure to learn optimal low q values is alleviated. This way, the values of q should increase, on average, approaching the 0.5 value. Differently, the proposed values exhibit a dependence on the group size that is conditioned on M. For mild group acceptance criteria (low M), having a big group of Responders is synonym of having a proposal easily accepted. In these circumstances, Proposers tend to offer less without risking having their proposals rejected, keeping this way more for themselves and exploiting the Responders. Oppositely, when groups agree upon stricter acceptance, having a big group of Responders means that more persons need to be convinced of the advantages of a proposal. This way, Proposers have to adapt, increase the offered values and sacrifice their share in order to have their proposals accepted. Contrarily to the sub-game perfect equilibrium played by fully rational agents, learning agents, under the proper group configurations, can indeed learn to be fair.

Acknowledgments

This research was supported by Fundação para a Ciência e Tecnologia (FCT) through grants SFRH/BD/94736/2013, PTDC/EEI-SII/5081/2014, PTDC/MAT/STA/3358/2014 and by multi-annual funding of CBMA and INESC-ID (under the projects UID/BIA/04050/2013 and UID/CEC/50021/2013 provided by FCT).

REFERENCES

[1] U. Fischbacher, C. M. Fong, and E. Fehr. Fairness, errors and the power of competition. *Journal of Economic Behavior & Organization*, 72(1):527–545, 2009.

[2] W. Güth, R. Schmittberger, and B. Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4):367–388, 1982.

[3] R. J. Kauffman, H. Lai, and C.-T. Ho. Incentive mechanisms, fairness and participation in online group-buying auctions. *Electronic Commerce Research and Applications*, 9(3):249–262, 2010.

[4] D. M. Messick, D. A. Moore, and M. H. Bazerman. Ultimatum bargaining with a group: Underestimating the importance of the decision rule. *Organizational Behavior and Human Decision Processes*, 69(2):87–101, 1997.

[5] M. J. Osborne. *An introduction to game theory*. Number 3. Oxford University Press New York, 2004.

[6] A. E. Roth and I. Erev. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8(1):164–212, 1995.

[7] F. P. Santos, F. C. Santos, A. Paiva, and J. M. Pacheco. Evolutionary dynamics of group fairness. *Journal of Theoretical Biology*, 378:96–102, 2015.