ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-1592

Execution Errors Enable the Evolution of Fairness in the Ultimatum Game

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

The goal of designing autonomous and successful agents is often attempted by providing mechanisms to choose actions that maximise some reward function. When agents interact with a static environment, the provided reward functions are well-defined and the implementation of traditional learning algorithms turns to be feasible. However, agents are not only intended to act in isolation. Often, they interact in dynamic multiagent systems whose decentralised nature of decision making, huge number of opponents and evolving behaviour stems a complex adaptive system [3]. This way, it is an important challenge to unveil the long term outcome of agents' strategies, both in terms of individual goals and social desirability [9]. This endeavour can be conveniently achieved through the employment of new tools from, e.g., population dynamics [4] and complex systems research, in order to grasp the effects of implementing agents whose strategies, even rational in the context of static environments, may turn to be disadvantageous (individually and socially) when successively applied by members of a dynamic population.

In this paper, we present a paradigmatic scenario in which behavioural errors are pernicious if committed in isolation, yet are the source of long-term success when considering adaptive populations. Moreover, errors support population states in which fairness (less inequality) is augmented. We assume that the goals and strategies of agents are formalised through the famous Ultimatum Game (**UG**) [2]. We focus on the changes regarding the frequency of agents adopting each strategy, over time. This process of social learning, essentially analogous to the evolution of animal traits in a population, enables us to use the tools of Evolutionary Game Theory (EGT), originally applied in the context of theoretical biology [4]. We describe analytically the behavioural outcome in a discretised strategy space of the UG, in the limit of small exploration rates (or the so-called mutations) [1]. This allows us to replicate the results of large-scale simulations [7], yet avoiding the burden of computational resources.

1 Ultimatum Game

The rules of UG are quite simple: some amount of a resource is conditionally endowed to one agent (Proposer) that must suggest a division with a Responder; secondly, the Responder will accept or reject the offer. The agents divide the money as it was proposed, if the Responder accepts. By rejecting, no one gets anything.

The rational behaviour in UG can be defined using a game-theoretical equilibrium analysis, through a simple backward induction. Facing the decision of rejecting (earn 0) or accepting (earn some money, even if a really small quantity), the Responder would always prefer to accept. Secure about this certain acceptance, the Proposer will offer the minimum possible, maximising his own share. Denoting by p the fraction of the resource offered by the Proposer, $p \in [0, 1]$, and by q the acceptance threshold of the Responder, $q \in [0, 1]$, acceptance will occur whenever $p \geq q$ and the *subgame perfect equilibrium* [5] of this game is defined by values of p and q slightly above 0. This outcome is said to be unfair, as it presents a profound inequality between the gains of Proposer and Responder. The strategies of agents that value fairness are characterised by prescribing a more equalitarian outcome: a fair Proposer suggests an offer close to 0.5 and a fair Responder rejects unfair offers, much lower than 0.5 (i.e. $p = 0.5$ and $q = 0.5$).

2 Analytical Framework and Results

We consider the existence of two populations (Proposers and Responders) each one composed by Z agents. Let us assume that a Proposer and a Responder may choose one of S strategies, corresponding to increasing divisions of 1. A Proposer choosing strategy $m \in \{1, 2, ..., S\}$ will offer the corresponding to $p_m = \frac{1}{S}m$ and a Responder choosing strategy $n \in \{1, 2, ..., S\}$ will accept any offer equal or above $q_n = \frac{1}{S}n$. The two-person encounter between a Proposer and a Responder thus yields $1 - p_m$ to the Proposers and p_m to the Responder if the proposal is accepted ($n : q_n \leq p_m$) and 0 to both agents otherwise.

We are concerned with the role of systematic errors in the execution of the desired strategy, namely, by the Responders. We assume that each Responder with strategy n (and threshold of acceptance q_n) will actually use a threshold of q'_n, calculated as $q'_n = q_n + U(-\epsilon, \epsilon)$, where $U(-\epsilon, \epsilon)$ corresponds to an error sampled from a uniform distribution between $-\epsilon$ and ϵ. Thereby, a Responder (using strategy n) accepts an offer $p_m \in [q_n - \epsilon, q_n + \epsilon]$ with a probability given by $P(q'_n \leq p_m) = P(q_n + U(-\epsilon, \epsilon) \leq p_m) = P(U(-\epsilon, \epsilon) \leq p_m - q_n) = \int_{-\epsilon}^{p_m - q_n} \frac{1}{2\epsilon} d(p_m - q_n) = \frac{p_m - q_n + \epsilon}{2\epsilon}$. The probability of acceptance is 0 if $p_m < q_n - \epsilon$ and is 1 if $p_m \geq q_n + \epsilon$. The resulting payoff of a pair (proposal, acceptance threshold) is, thereby, linearly weighted by the probability of acceptance, considering the execution error (ϵ).

As we assume a well-mixed population, the fitness of each individual is given by the average payoff earned when playing with all the agents in the opposite population. Considering the existence of S different strategies in the opposite population of the one from which agent A belongs; denoting k_i as the number of agents using strategy i, in the opposite population of A; and regarding $R_{j,i}$ as the payoff (reward) earned by an agent A using strategy j, against an agent with

[1] INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, email: fernando.pedro@tecnico.ulisboa.pt
[2] ATP-Group, Lisboa, Portugal
[3] CBMA and Departamento de Matemática e Aplicações, Universidade do Minho

strategy i, the fitness of agent A is given by $F_{A_j} = \sum_{i=1}^{S} \frac{k_i}{Z} R_{j,i}$.

The adoption of strategies will evolve following an imitation process. We assume that at each step two agents are chosen, one that will imitate (agent A) and one whose fitness and strategy will serve as model (agent B). The imitation probability will be calculated using a function — $(1+e^{-\beta(F_B-F_A)})^{-1}$ — that grows monotonously with the fitness difference $F_B - F_A$ [10]. The variable β in the equation above is well-suited to control the selection pressure, allowing to manipulate the extent to which imitation depends on the fitness difference. Assuming that two agents are randomly sampled from the population in which k_i agents are using strategy i (the remaining are using strategy j), the probability of having ± 1 (more or less 1) individual using strategy i is given by

$$T^{\pm}(k_i) = \frac{Z-k_i}{Z}\frac{k_i}{Z-1}\left(1+e^{\mp\beta(F_i(\overline{k}_s)-F_j(\overline{k}_s))}\right)^{-1} \qquad (1)$$

assuming that in the opposite population the number of agents using another strategy s is $\overline{k}_s$ and that the population size is Z. Note that $\frac{Z-k_i}{Z}$ (and $\frac{k_i}{Z-1}$) represent the sampling probabilities of choosing one agent with strategy $j(i)$.

With probability μ, a mutation occurs and individuals change their strategy to a random one, exploring a new behaviour regardless the observation of others. The imitation process described above will occur with probability $(1-\mu)$. If we assume that $\mu \to 0$, we are able to derive analytical conclusions through a simpler apparatus [1]. Under this regime in which mutations are extremely rare, a *mutant* strategy will either fixate in the population or will completely vanish [1] and the number of different strategies present in the population is at most two. The time between two mutation events is usually so large that the population will always evolve to a monomorphic state (i.e., all agents using the same strategy) before the next mutation occurs. This fact allows us to conveniently use Equation (1) in the calculation of the transition probabilities between intermediate states (where a mutant is still invading and two strategies co-exist in the population). The transitions between monomorphic states are described through the fixation probability of every single mutant of strategy i in every resident population of strategy j. A strategy i will fixate in a population composed by $Z-1$ individuals using strategy j with a probability given by

$$\rho_{i\to j}(\overline{k}_s) = \frac{1-e^{-\beta\Delta F(\overline{k}_s)}}{1-e^{-Z\beta\Delta F(\overline{k}_s)}} \qquad (2)$$

These probabilities define an embedded Markov Chain, governed by the stochastic matrix T, in which $T_{i,j} = \rho_{i\to j}$ defines the fixation probability of a mutant with strategy i in a population with $Z-1$ individuals using strategy j. A derivation from Equation (1) to Equation (2) and more details regarding Equation (2) can be found in [4, 10]. To calculate π, the stationary distribution of this Markov Process, we compute the normalised eigenvector associated with the eigenvalue 1 of the transposed of T. $\pi_{a,b}$ represents the fraction of time, on average, that is spent when the population of Proposers is using strategy a and the population of Responders is using strategy b. The number of possible states depends on the discretisation considered in the strategy space available in Ultimatum Game. If the Proposer and Responder have, each, S available strategies, there are S^2 different monomorphic states. The resulting average fitness is provided by the average fitness of a population in each monomorphic state, weighted by the time spent in that state. Thereby, the average fitness of the population of Proposers is given by $\overline{F} = \sum_{a=1,b=1}^{S} \pi_{a,b} R_{a,b}$

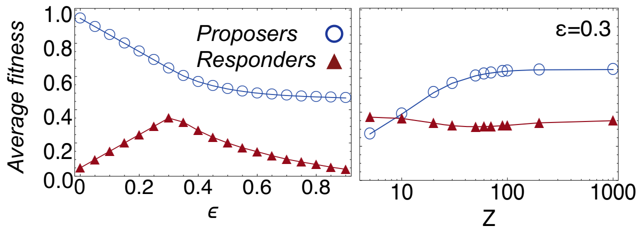

Figure 1. Analytical results showing the role of errors (ϵ) and population sizes (Z) in the overall fitness of Proposers and Responders. $\beta = 10$, $Z = 100$, $S = 20$

and the average fitness of the population of Responders is given by $\overline{F} = \sum_{a=1,b=1}^{S} \pi_{a,b} R_{b,a}$.

With this framework, we are able to show that the fitness of the Responders will be maximised if they commit a significant execution error, sampled from an interval close to $[-0.3, 0.3]$ (Figure 1). Moreover, this framework is well suited to capture the stochastic effects of considering finite populations and even the role of different population sizes. We additionally verified that increasing β and Z promotes determinism in the imitation process which benefits the fitness of the Proposers and undermines the fair distribution of gains between Proposers and Responders. Even employing a different methodology, these results are in line with the discussions performed in [6, 8].

ACKNOWLEDGEMENTS

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] Drew Fudenberg and Lorens A Imhof, 'Imitation processes with small mutations', *J Econ Theory*, **131**(1), 251–262, (2006).
[2] Werner Güth, Rolf Schmittberger, and Bernd Schwarze, 'An experimental analysis of ultimatum bargaining', *J Econ Behav Organ*, **3**(4), 367–388, (1982).
[3] John H Miller and Scott E Page, *Complex adaptive systems: an introduction to computational models of social life*, Princeton University Press, 2009.
[4] Martin A Nowak, *Evolutionary dynamics: exploring the equations of life*, Harvard University Press, 2006.
[5] Martin J Osborne, *An introduction to game theory*, Oxford University Press New York, 2004.
[6] David G Rand, Corina E Tarnita, Hisashi Ohtsuki, and Martin A Nowak, 'Evolution of fairness in the one-shot anonymous ultimatum game', *Proc Natl Acad Sci USA*, **110**(7), 2581–2586, (2013).
[7] Fernando P Santos, Francisco C Santos, and Ana Paiva, 'The evolutionary perks of being irrational', in *Proceedings of the 2015 AAMAS*, pp. 1847–1848. International Foundation for Autonomous Agents and Multiagent Systems, (2015).
[8] Fernando P Santos, Francisco C Santos, Ana Paiva, and Jorge M Pacheco, 'Evolutionary dynamics of group fairness', *J Theor Biol*, **378**, 96–102, (2015).
[9] Yoav Shoham, Rob Powers, and Trond Grenager, 'If multi-agent learning is the answer, what is the question?', *Artif Intell*, **171**(7), 365–377, (2007).
[10] Arne Traulsen, Martin A Nowak, and Jorge M Pacheco, 'Stochastic dynamics of invasion and fixation', *Phys Rev E*, **74**(1), 011909, (2006).