The Evolutionary Perks of Being Irrational

(Extended Abstract)

Fernando P. Santos
INESC-ID and Instituto Superior Técnico, Universidade de Lisboa
Taguspark, Av. Prof. Cavaco Silva
2780-990 Porto Salvo, Portugal
fernando.pedro@ist.utl.pt

Francisco C. Santos
INESC-ID and Instituto Superior Técnico, Universidade de Lisboa
Taguspark, Av. Prof. Cavaco Silva
2780-990 Porto Salvo, Portugal
franciscocsantos@ist.utl.pt

Ana Paiva
INESC-ID and Instituto Superior Técnico, Universidade de Lisboa
Taguspark, Av. Prof. Cavaco Silva
2780-990 Porto Salvo, Portugal
ana.paiva@inesc-id.pt

ABSTRACT

Typical implementations of artificial intelligent agents suggest that actions should be chosen in order to maximise some reward function. This naturally complies the philosophy behind rational choice theory. Yet, this heuristic may not always provide long-term success to the considered agents. In this paper, we stress the need to consider the self-organised and frequency-dependent nature of the environment, when designing agents that act in complex adaptive systems. We resort to the tools of evolutionary game theory, combined with a paradigmatic scenario of a population of self-regarding agents playing Ultimatum Game, to describe the dynamical impact of individual mistakes on collective behaviour. We resort to agent based simulations to show that that seemingly disadvantageous and irrational errors become the source of individual and collective long-term success.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Distributed Artificial Intelligence-Multiagent systems

General Terms

Economics, Algorithms, Theory

Keywords

Complex systems, Evolution and co-evolution, Emergent behaviour, Simulation techniques, Ultimatum Game, Biologically-inspired approaches and methods

1. INTRODUCTION

The goals of artificial agents are often formalised through the notion of utility, and the capability to choose actions intelligently is fulfilled through utility-maximisation heuristics. Thereby, the framework of rational choice theory embodies itself the notion of intelligence as *rationality* may conveniently ground stylised models of human behaviour. When empirical evidence shows that humans systematically deviate from the rational model, explanations often suggest the lack of information or computational power. Consequently, the concept of bounded rationality relaxes the strongest assumptions of the pure rational model [5, 4]. Either way, the existence of seemingly irrational decisions is often reported as disadvantageous.

In this paper, we present a paradigmatic scenario in which irrational behaviour may be the source of long-term success, if we consider a complex system composed by self-regarding agents. We assume that the goals and strategies of agents are formalised through the famous Ultimatum Game (**UG**) [2]. The rules of this game are simple: two players interact in two distinct roles. One is called the *Proposer* and the other is denominated *Responder*. The game is composed by two subgames, one played by each role. First, some amount of a given resource, e.g. money, is conditionally endowed to the Proposer, and this agent must then suggest a division with the 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, none of them will get anything. The rational behaviour in UG can be defined through the notion of *subgame perfect equilibrium*. If one divides the UG in two stages, as suggested above, it is possible to apply the method of backward induction. The last agent to play is the Responder. Facing the decision of rejecting (earn 0) or accepting (earn some money, even if a really small quantity), this agent would rationally prefer to Accept. Secure about this deterministic acceptance, the Proposer will offer the minimum possible, maximising his own share. Thereby, the sub-game perfect equilibrium predicts offers close to the minimum by the Proposers and unconditional acceptance by the Responders. Intriguingly however, there is a myriad of studies that account for an irrational behaviour by human beings when playing this game [2, 1].

We model a finite population of adaptive agents that coevolve by imitating the best observed actions [6]. 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, enable us to use tools of Evolutionary Game Theory (EGT), originally applied in the context of theoret-

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

ical biology [3]. We compute the behavioural outcome in a set of agent based simulations detailed in the next section.

2. AGENT BASED SIMULATIONS

We assume that agents may choose a continuity of strategies. In the population of the Proposers, the strategy of each agent is characterised by value $p \in [0, 1]$, the offer being proposed. In the population of Responders, each individual action is defined by $q \in [0, 1]$, the threshold of acceptance. A proposal is accepted if $p \geq q$. In this case, the Proposer earns $1 - p$ and the Responder earns p. If $p < q$ both agents earn 0. We employ a general procedure to simulate evolving agents in the context of EGT. At each time step, a population is randomly picked, from which two agents are chosen (agents A and B). The fitness of each agent (f_A and f_B) is calculated by averaging their returns when interacting with all agent from the opposite population. Agent A will then imitate agent B with a probability provided by a sigmoid function – $\left(1 + e^{\beta(f_A - f_B)}\right)^{-1}$ [6]. With a small probability of μ, imitation will not take place and agent A updates the own strategy to a randomly picked one, between 0 and 1. In biology, this corresponds to a genetic mutation while, in social learning and cultural evolution, this mimics the random exploration of behaviours. A similar procedure may be found, for instance, in reinforcement learning with the so-called ϵ-greedy methods. The variable β in the equation above is well-suited to control the selection strength, allowing to manipulate the extent to which imitation depends on the fitness difference. The same procedure takes place in the opposite population. When $2Z$ steps of imitation occur, Z in each population, we say that one generation has elapsed. We evolve our system for 5000 generations, and we save the average fitness and average strategy used, for each population. Agents start with random strategies, sampled from a uniform distribution between 0 and 1. We repeat the simulation for 50 times, each time starting with random conditions. The results presented (average fitness) correspond to a time average over all generations and an ensemble average over all repetitions. In all plays done by the Responders, a noise factor, translating a systematic behavioural error, will be added their base strategy. Thus, the real action employed by Responders will correspond to q', their base strategy (q), plus $U(-\epsilon, \epsilon)$, a random value between $-\epsilon$ and ϵ sampled from a uniform distribution. Naturally, the values of q' are truncated to remain in the closed domain $[0, 1]$.

3. RESULTS

We show that the fitness of the Responders will be maximised if they commit an execution error sampled from an interval close to $[-0.3, 0.3]$. In Figure 1, we show how the average fitness of Proposers (circles) and Responders (triangles) is affected by changing the range of possible execution errors (ϵ) committed by Responders. If the error increases, Responders are endowed with increased fitness. The Proposers are always harmed by the erroneous behaviour of Responders. The *sub-game perfect equilibrium* prediction poses that Proposers will earn all the pot when they offer almost nothing to the Responder, assuming an unconditional acceptance by this agent. Yet, if it is assumed that Responders will commit execution errors, which, in the case of heighten the threshold of acceptance may be seen as an irrational

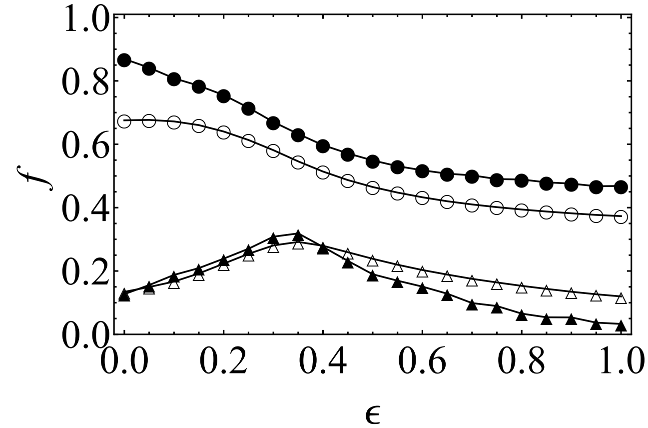

Figure 1: The effects of varying the span of error (ϵ) committed by the Responders in the average fitness of Proposers (circles) and average fitness of Responders (triangles), for $\mu = 0.001$ (filled) and $\mu = 0.1$ (empty). It is beneficial for the Responder to behave erroneously, to some extent. If Responders reject irrationally some proposals, the Proposers have to adapt and start to offer more, beneficing, in the long run, the Responders; If Responders reject too much, they will harm themselves and the population of Proposers. As verified, these conclusions remain valid for a wide range of mutation rates (μ). Other parameters: $Z = 100$, $\beta = 10$.

behaviour, the Proposers necessarily have to adapt to have their proposals accepted and earn some payoff. This adaptation leads to an increase of the offer, favouring the average fitness of the Responders. Additionally, we note that if the Responders error unreasonably, both Proposers and Responders will be impaired. Overall, we provide a minimal model showing how erroneous (or irrational) moves can be a rewarding option at the population level, while fostering fairness in pairwise endeavours.

4. ACKNOWLEDGMENTS

This research was supported by FCT-Portugal through grants SFRH/BD/94736/2013, EXPL/EEI-SII/2556/2013, and UID/CEC/50021/2013.

REFERENCES

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

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

[3] J. Maynard-Smith and G. Price. The logic of animal conflict. *Nature*, 246:15, 1973.

[4] S. J. Russell. Rationality and intelligence. *Artificial Intelligence*, 94(1):57–77, 1997.

[5] H. A. Simon. Bounded rationality and organizational learning. *Organization Science*, 2(1):125–134, 1991.

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