The "Favors Game": A Framework to Study the Emergence of Cooperation through Social Importance

(Extended Abstract)

Pedro Sequeira, Samuel Mascarenhas, Francisco S. Melo, and Ana Paiva
INESC-ID and Instituto Superior Técnico, University of Lisbon
Av. Prof. Dr. Cavaco Silva, 2744-016 Porto Salvo, Portugal
pedro.sequeira@gaips.inesc-id.pt

ABSTRACT

Several agent-based frameworks have been proposed to investigate the possible reasons that lead humans to act in the interest of others while giving up individual gains. In this paper we propose a novel framework for analyzing this phenomenon based on the notions of social importance and local discrimination. We propose a "favors game", where a recipient agent can "claim" a favor to a donor agent, which may in turn "confer" its request at the expense of a certain cost. The proposed framework allows us to study the conditions under which cooperation occurs and the dynamics of the relationships formed within a population.

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Economics, Human Factors, Theory

Keywords

social importance, favors game, cooperation, social agents

1. INTRODUCTION

Cooperation appears in nature as an organizational mechanism capable of enhancing the reproductive power of individuals that live in groups and help each other. The subject who cooperates, or *donor*, pays a cost so that the fitness of another individual, the *recipient*, is increased [1, 5]. Having in mind that individuals interact with each other many times throughout their lives, the success behind cooperation is that the instant cost inflicted by an altruistic act can be compensated by long-term benefits if others also help the donor in the future. However, in mixed populations, defectors will perform better on average than cooperators and natural selection based on competition will decrease the proportion of cooperators until they become extinct [1, 5, 9, 2].

Recently, a model was proposed with the goal of simulating the general dynamics of human social interaction [4].

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.*

The model is based on Kemper's *sociological theory of human motivation* [3], which postulates that we act in favor of those to whom we confer enough status and, conversely, the amount of collaboration we can claim from others depends on how much status we assume to have in their minds. This concept was modeled as *social importance* (SI), according to which individuals perform *claims* and *conferrals* when they believe they are sufficiently important to others. In this paper we propose a novel framework to analyze the emergence of cooperation between self-interested individuals. As a testbed, we developed a "favors game" where one agent may choose to claim or not a favor, and another agent may confer or not to it.

2. THE "FAVORS GAME"

In our framework, the notion of SI is translated into a scalar in $[0;1]$, denoted by $\sigma(a,b)$, asserting the extent to which individual a will voluntarily comply with the needs or desires of individual b. In other words, we can say that $\sigma(a,b)$ represents the SI that a attributes to b. Similarly, an individual can estimate its relevance in the perspective of others as a result of interactions. We thus denote by $\hat{\sigma}(b,a)$ the SI that individual a believes that b attributes to him.

2.1 Game Dynamics and Payoffs

We designed a simple game played by two individuals, which we refer to as the "favors game", that involves the exchange of work favors. We incorporated the ideas from the SI model by letting an individual play the role of *claimer* and choose whether to ask another individual, playing the role of *conferrer*, to help him perform some work that needs to be done. If the claimer asks the favor, the conferrer may concede to do it. If this is the case, the individuals will work together and both will incur in some cost, denoted by c_t. The claimer will receive the benefit of the work, which we denote by b_t. The conferrer may however decide not to confer to the claim, in which case the claimer will perform the work alone at the expense of some cost, denoted by c_a, and receive a benefit denoted by b_a. In this case, the conferrer does not incur in any expense nor he receives any benefit. We assume that some effort is associated with claiming a favor, and in our game this is reflected by an extra cost, denoted by c_c, to the player making the claim. The claimer may in alternative decide not to ask the favor, in which case two situations may occur: the conferrer may still decide to help by performing his work for him, the claimer thereby receiving all the work benefit b_a and the conferrer

Table 1: Payoffs for each player within the "favors game", where CL and $\overline{CL}$ refer to the actions of claiming and not claiming, respectively, and CO and $\overline{CO}$ refer to the actions of conferring and not conferring, respectively.

(a) Claimer payoff.

	CO	$\overline{CO}$
CL	1	−0.5
$\overline{CL}$	1	0

(b) Conferrer payoff.

	CL	$\overline{CL}$
CO	−0.5	−1
$\overline{CO}$	0	0

all the work cost c_a;[1] otherwise the claimer performs the work alone, receiving the benefit b_a and paying the cost c_a.

To study the emergence of cooperation in competitive contexts we may set the values for the benefits and costs so that it best fits our purpose. For example, if we set $c_c = -0.5$, $c_a = -1$, $b_a = 1$, $c_t = -0.5$, $b_t = 2$, yielding the payoff matrices for each player depicted in Table 1, we force mutual defection to be the *dominant* action. By doing so, and without an external mechanism, natural selection based on competition will promote the emergence of defection, making the "favors game" a good testbed for the study of the dynamics behind the SI mechanism.

2.2 Claiming and Conferring

A claim is made whenever an individual "feels worthy" of performing such request, according to the perceived state of his relationship with the conferrer [4]. In our framework this "worthiness" is reflected by the probability $\mathbb{P}_{cl}(a,b)$ of individual a asking individual b a favor given by

$$\mathbb{P}_{cl}(a,b) = \mathcal{B}(\hat{\sigma}(b,a), \theta_{cl}(a)), \tag{1}$$

where $\mathcal{B}(x,x_0) = (1 + e^{-\beta(x-x_0)})^{-1}$. $\mathcal{B}(x,x_0)$ thus corresponds to the Boltzmann function centered around x_0, with inverse temperature β determining the sharpness of the transition from $\mathcal{B} = 0$ to $\mathcal{B} = 1$. The parameter $\theta_{cl}(a)$ is a *claim threshold* taking values in $[0,1]$ that represents the *predisposition* of individual a for asking others for favors. As can be seen from (1), the higher the threshold $\theta_{cl}(a)$ the harder it will be for a to perform claims. Also, the greater a's belief regarding the quality of his relationship with b, given by $\hat{\sigma}(b,a)$, the more likely it is to ask b a favor.

We model the "favors game" as a two-stage game: claims occur in the first stage of the game, conferrals in the second. As such, we must also define two probabilities and thresholds associated with conferring after someone made a claim, which we refer to as an *explicit conferral*, and when a claimer did not ask for help, corresponding to an *implicit conferral*. Following a similar approach to the claim action, we define the probability of an individual a performing respectively an explicit and an implicit conferral to b by:

$$\mathbb{P}_{ce}(a,b) = \mathcal{B}(\sigma(a,b), \theta_{ce}(a)), \tag{2}$$

$$\mathbb{P}_{ci}(a,b) = \mathcal{B}(\sigma(a,b), \theta_{ci}(a)), \tag{3}$$

where $\theta_{ce}(a)$ and $\theta_{ci}(a)$ are the explicit and implicit thresholds, respectively, and represent the predisposition of individual a to perform conferrals. The general game strategy of an individual playing the "favors game" is thus determined by the values of these three thresholds. Formally, we denote the strategy of a player as the vector $\Theta = [\theta_{cl}, \theta_{ce}, \theta_{ci}]$.

[1] We assume that because the claimer did not ask the conferrer a favor, the latter will perform the work by himself.

3. CONCLUSIONS

In this paper we proposed a framework inspired by a theory that simulates the general dynamics of human social interaction through the concept of social importance. We propose an inherently competitive "favors game" where agents "claim" and "confer" favors, making it a good testbed to study the emergence of cooperation between self-interested interacting individuals. Our "favors game" is similar to the "repeated helping game" [7, 8] studied in the context of indirect reciprocity involving the exchange of "help" between donors and recipients. In the "favors game" however, the decision on whether to help another individual is contingent on one agent's social predisposition and on previous interactions with that specific individual. In addition, our matching mechanism based on the SI levels resembles that of network reciprocity [6], in which the matching of interacting players is based on a spatial structure of the population.

The SI model in [4] is a dynamic mechanism that changes through time as a consequence of repeated interactions. In the future we intend to study the dynamics of the proposed SI framework in the context of the "favors game" by performing several agent-based computational simulations and study the conditions under which cooperation occurs. We also intend to analyze the dynamics of the relationships formed within a population by performing sensitivity analysis on the social factors underlying the proposed mechanism.

Acknowledgments

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013. The first author acknowledges a BDP grant CMUP-ERI/HCI/0051/2013 from the INSIDE project funded by the FCT. The second author acknowledges the PhD grant SFRH/BD/62174/2009 from the FCT.

REFERENCES

[1] R. M. Axelrod. *The evolution of cooperation*. Basic Books, Inc., Publishers, New York, USA, 1984.

[2] R. Dawkins. *The Selfish Gene*. Oxford University Press, New York, USA, 30th anniv. ed. edition, 2006.

[3] T. D. Kemper. *Status, Power and Ritual Interaction*. Ashgate, 2011.

[4] S. Mascarenhas, R. Prada, A. Paiva, and G. J. Hofstede. Social importance dynamics: A model for culturally-adaptive agents. In *Int. Virt. Agents*, number 8108, pages 325–338. Springer Berlin Heidelberg, 2013.

[5] M. A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, Dec. 2006.

[6] M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, 10 1992.

[7] M. A. Nowak and K. Sigmund. The dynamics of indirect reciprocity. *Journal of Theoretical Biology*, pages 561–574, 1998.

[8] I. Seinen and A. Schram. Social status and group norms: Indirect reciprocity in a repeated helping experiment. *Europ. Econ. Review*, 50(3):581–602, 2006.

[9] R. Trivers. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1):35–57, 1971.