Selfishness Drives Collective Cooperation and Network Formation

Korosh Mahmoodi and Cleotilde Gonzalez, Dynamic Decision Making Laboratory, Department of Social and Decision Sciences, Carnegie Mellon University

1. INTRODUCTION

Emergence of collective cooperation in a society composed of selfish individuals is known as the *altruism paradox*, and it has preoccupied biologists, sociologists, and cognitive scientists alike for centuries. This paradox has been commonly studied in the well-known Prisoner's Dilemma (PD). PD is a common abstraction of essential elements of many naturalistic situations involving the trade-off between cooperative and selfish behavior [Rapoport and Chammah 1965; Nowak and Sigmund 1993; Gonzalez et al. 2015]. When two players cooperate, each of them gains the payoff R, and when both players defect, each of them has a penalty P. If the player i defects and player j cooperates, player i gains the payoff T and player j gains the payoff S and vice versa. The constraints on the payoff values are $T > R > P > S$ and $S + T < 2R$. The temptation to defect is established by $T > R$. The dilemma is that each individual would prefer to defect (defection provides a higher reward to the individual) and consequently a pair of agents will end up in a $D - D$ situation with the minimum payoff of $2P$. How do then individuals learn that cooperation is mutually beneficial in the long-term?

A popular answer to this question, known as *network reciprocity*, was introduced by Nowak and May [1992]. They placed 400×400 Agents in a two-dimensional spatial array where each agent plays the PD with 8 neighbors and act by copying the strategy of their richest neighbor (including itself), making a decision by comparing the other's payoff to the agent's own payoff. Network reciprocity relies on the assumptions of existence of a network structure and on the awareness of the payoffs of interconnected agents. In reality, some experiments suggest that humans do not consider others' payoffs when making their decisions [Fischbacher et al. 2001], and others suggest that although humans are not consistently selfish they adjust their level of "selfishness" dynamically according to the other player's actions [Gonzalez et al. 2015]. Furthermore, it is unclear how cooperation emerges in situations in which there is no network of agents in the first place.

In this paper we explore the emergence of collective cooperation from selfish individuals, in situations in which there is no explicit a prior network structure (i.e., no predefined connections and no knowledge of others' outcomes). We present simulation results from a new algorithm (i.e., *Living Thing*, LT). LT relies on two novel mechanisms: Selfish Imitation (SI) and Selfish Attachment (SA). SI is a decision motivated by an agent's desire to increase its own benefit: agent i, playing with agent j, increases the chance of imitating agent j's strategy if after playing with agent j its payoff is higher than its own previous payoff. Similarly, SA is a decision of agent i to connect with agent j. The probability of this connection increases if after playing with agent j, i observes that its payoff is higher than its own previous payoff. In other words, in LT improvement of the agent's own payoff is the only learning incentive. Results from simulations demonstrate that cooperation can emerge and survive between agents acting exclusively on their own individual benefit and without the existence of any predefined network. Furthermore, we show that a dynamic network emerges from these selfish interactions.

2. LIVING THING (LT) ALGORITHM AND EMERGENCE OF COOPERATION FROM SELFISHNESS

The LT algorithm originally presented in [Mahmoodi and Gonzalez 2019], proposes the following steps executed in each time step (trial). First, each agent of a randomly picked pair of agents, which agreed to play with each other, makes its potential decision to play C or D in the PD based on reinforcement of past actions (RL). Then an agent may decide to stick with its own decision (from RL) or to copy the decision of its partner based on its experience of imitating it (SI). After making a decision based on RL or SI, the agent receives the outcome according to the PD payoff matrix. When the agent used RL, and if playing C (D) increased its payoff relative to the previous one, then the agent increases the chance of playing C (D) next time it pairs up with the same agent. If the agent used SI instead, and after playing with its partner received a payoff higher than its previous payoff, then it will increase the chance to imitate this partner next time they pair up. If playing with a partner increased the benefit of the agent relative to its previous payoff then the agent will increase the chances of playing with this partner again in the future (SA).

We studied a system with $N = 100$ agents interacting over the course of 10^7 trials. Initially all the agents are defectors, have payoff of zero, have a high chance to stay as defector (99/100), have more chance to use RL over SI (99/100), and have equal chance to pair up with other agents (1/(N-1)). These chances (to pick one of the choices) update at every trial for the two playing agents from $A/(A + B)$ and $B/(A + B)$ to $(A \pm \Delta)/(A + B)$ and $(B \mp \Delta)/(A + B)$ if selecting the choice corresponding to A (for example playing C) increased (decreased) its payoff relative to the previous payoff. We set $\Delta = 10$. Δ shows the sensitivity of the agent to the feedback from its two last payoffs. Smaller Δ decreases the rate of reaching to cooperation but doesn't change the dynamic properties of the system. $R = 1$, $P = 0$ and $S = 0$ so, the maximal possible value of T is 2. We selected the value $T = 1.9$, which gives a very strong incentive to cheat.

Panels A and B of Figure 1 show the proportion of cooperators of a model that uses RL only compared to a model using SI (Panel A); and the same proportions of the model when SA is added (Panel B). Note that the proportion of cooperators with RL only fluctuates around 0.3, regardless of SA being added (compare Panel A to Panel B, blue curves). In contrast, the proportion of cooperators with SI fluctuates around 0.9 (Panel A, orange curve). When SA is added to SI, the proportion of cooperators increases with less fluctuations (Panel B, orange curve). Panels C and D show the probability of imitation of one agent and other 9 agents in a system with 10 agents. Panel C demonstrates the increase in the chances of imitating other agents in a system with SI. Panel D demonstrates how an agent prefers to imitate some of the agents more often than others in a system with SI and SA.

3. NETWORK FORMATION THROUGH SELFISH ATTACHMENT

Figure 2a demonstrates the chance of a random agent (represented by a dot in the center) pairing with the other 99 agents at two different times: $t = 10^2$ (left panel) and 10^6 (right panel). The thickness of the lines represent the chance of pairing of the agent in the center with the other 99 agents at two times in the simulation (the thicker line represents a higher probability), when the agents used SI + SA to update. At $t = 10^2$, we observe an almost uniform distribution of the connections of an agent to the others (left panel), but later, the agent learned to prefer to stick to some partners for most of the time (right panel).

Figure 2b shows the histogram of the probability of the pairings for all the agents in Figure 2 (right panel).This histogram, plotted in a log-log scale, shows an inverse power law distribution $\propto 1/P^{\beta}$ with complexity index of $\beta = 1.3$, rather than having a Poisson distribution, suggesting that the emergence of this network is not stochastic and is dynamic and complex.

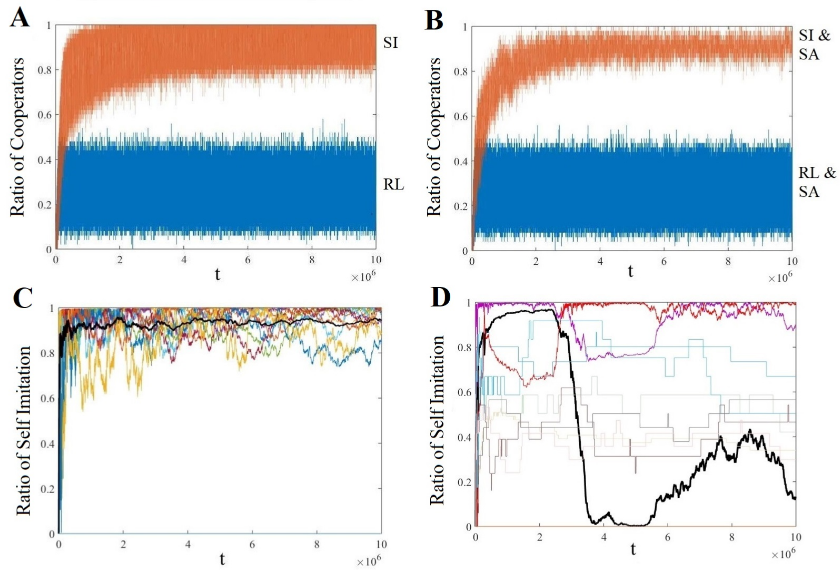

Fig. 1. Panels A and B: Proportion of cooperators (out of $N = 100$). A: Blue curve shows the proportion of cooperators using RL only and orange curve are those using SI. Panel B: Blue curve shows RL+SA and orange curve shows SA+SI. Panels C and D: Emergence and evolution of imitation chances of an agent the other 9 agents in a system with $N = 10$. Panel C: uses SI. The thicker, black curve is the average of all the nine imitation chances. Panel D: uses SI + SA. The thicker, black curve is the chance of an agent to play with other agents, which emulates the purple curve of imitating the same agent.

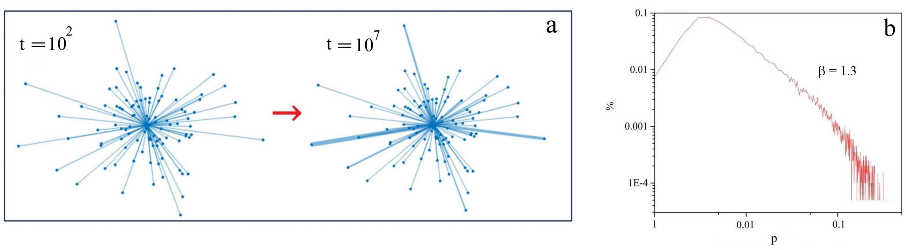

Fig. 2. Panel a: each dot represents an agent. The dot in the center shows the agent of interest and the other 99 dots are connected to it with lines. The thickness of each line is proportional to the chance of the corresponding agents to pair up at time $t = 10^2$ (left figure) and $t = 10^7$ (right figure). Panel b: The histogram of the chances for pairing up between all possible pairs at $t = 10^7$, in a log log scale.

4. DISCUSSION

We present the LT algorithm to resolve the altruism paradox in the absence of an explicit network structure and in the absence of explicit awareness of others outcomes. A main component of the algorithm is SI which lets the agent adapt strategies of others if doing so increases its own payoff relative to its last one. SI lets selfish agents realize the benefit of cooperation. Adding SA to SI lets each agent learn which agents help increase its own payoff with respect to its last one. The combination of SI and SA generates a dynamic network that demonstrates to be complex with an inverse power law distribution of pairing chances.

REFERENCES

Urs Fischbacher, Simon Gächter, and Ernst Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics letters* 71, 3 (2001), 397–404.

Cleotilde Gonzalez, Noam Ben-Asher, Jolie M. Martin, and Varun Dutt. 2015. A Cognitive Model of Dynamic Cooperation With Varied Interdependency Information. *Cognitive Science* 39, 3 (2015), 457–495. DOI:http://dx.doi.org/10.1111/cogs.12170

Korosh Mahmoodi and Cleotilde Gonzalez. 2019. Emergence of Collective Cooperation from Selfish-Imitation and Selfish-Attachment. *Submitted to the cognitive science conference* (2019).

Martin Nowak and Karl Sigmund. 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364, 6432 (1993), 56–58.

Martin A Nowak and Robert M May. 1992. Evolutionary games and spatial chaos. *Nature* 359, 6398 (1992), 826.

A. Rapoport and A. M. Chammah. 1965. *Prisoner's Dilemma: A Study in Conflict and Cooperation*. University of Michigan Press, Ann Arbor.