Cost-efficiency of institutional reward and punishment in cooperation dilemmas

Duong, Manh Hong; Han, The Anh

DOI:
10.1162/isal_a_00517

Cost-efficiency of institutional reward and punishment in cooperation dilemmas

Manh Hong Duong[1] and The Anh Han[2]

[1] School of Mathematics, University of Birmingham, B15 2TT, UK. Email: h.duong@bham.ac.uk
[2] School of Computing, Media and Art, Teeside University, TS1 3BX, UK. Email: T.Han@tees.ac.uk

Introduction. A central challenge in biological, computational and social sciences is to understand the evolution of cooperation within populations of self-regarding individuals and mechanisms that promote it (Perc et al., 2017; Yang et al., 2018; Han, 2013). To this extent, various mechanisms have been revealed and studied using methods from evolutionary game theory, statistical physics and agent-based modelling and simulations (Maynard-Smith, 1982; Hofbauer and Sigmund, 1998; Perc et al., 2017). They include both endogenous and exogenous mechanisms such as kin and group selection, direct and indirect reciprocity, spatial networks (Nowak, 2006b), reward and punishment (Sigmund et al., 2001), and pre-commitments (Han et al., 2015). Institutional incentives, positive (reward) and negative (punishment), are among of the most important ones (Sigmund et al., 2001; Van Lange et al., 2014). In institutional incentives, an external decision maker (e.g. institutions such as the United Nations and the European Union) who has a budget to interfere in the population in order to achieve a desirable outcome, for instance to ensure a desired level of cooperation. Providing incentives for promoting cooperation is costly and it is thus important to optimize the cost while ensuring a sufficient level of cooperation (Ostrom, 1990; Chen et al., 2015; Cimpeanu et al., 2021). In the literature, evolution of populations can be studied using either a deterministic approach, which utilizes the continuous replicator dynamics assuming infinite populations, or a stochastic approach, which employs Markov chain for modelling finite populations. For infinite populations, Wang et al. (2019) has recently exploited optimal control theory to provide an analytical solution for cost optimization of institutional incentives. This work therefore does not take into account various stochastic effects of evolutionary dynamics such as mutation and those resulting from behavioural update (Nowak et al., 2004). This might be problematic since undesired behaviours can reoccur over time, via mutation or when incentives were not strong or effective enough in the past. Moreover, a key factor in behavioural update, the intensity of selection (Sigmund, 2010)—which determines how strongly an individual bases her decision to copy another individual's strategy on fitness difference and is absent in the continuous approach—might influence the incentivisation strategy and its cost efficiency as well. For finite populations, so far this problem has been investigated primarily based on agent-based and numerical simulations (Sasaki et al., 2012; Han and Tran-Thanh, 2018; Cimpeanu et al., 2019). In this extended abstract, starting from a finite population framework in (Han and Tran-Thanh, 2018), we summarize a recent publication (Duong and Han, 2021) that provides a *rigorous analysis*, supported by numerical simulations, for this problem and discuss open problems in this emerging research direction.

Models and Methods. We consider a well-mixed, finite population of N self-regarding individuals or players, who interact with each other using one of the following cooperation dilemmas, namely the Donation Game (DG) and the Public Goods Game (PGG). We adopt here the finite population dynamics with the Fermi strategy update rule (Traulsen et al., 2006), stating that a player A with fitness f_A adopts the strategy of another player B with fitness f_B with a probability given by, $P_{A,B} = \left(1 + e^{-\beta(f_B - f_A)}\right)^{-1}$, where β represents the intensity of selection. To reward a cooperator (resp., punish a defector), the institution has to pay (fine) an amount θ (resp., θ) so that the cooperator's (defector's) payoff increases (decreases) by θ. The population dynamics are modelled using an absorbing Markov chain consisting of $(N+1)$ states, $\{S_0, ..., S_N\}$, where S_i represents a population with i C players. S_0 and S_N are absorbing states. Let $U = \{u_{ij}\}_{i,j=1}^{N-1}$ denote the transition matrix between the $N-1$ transient states, $\{S_1, ..., S_{N-1}\}$. The transition probabilities can be defined as follows, for $1 \leq i \leq N-1$:

$$u_{i,i\pm j} = 0 \qquad \text{for all } j \geq 2,$$
$$u_{i,i\pm 1} = \frac{N-i}{N}\frac{i}{N}\left(1 + e^{\mp\beta[\Pi_C(i) - \Pi_D(i) + \theta]}\right)^{-1},$$
$$u_{i,i} = 1 - u_{i,i+1} - u_{i,i-1},$$

where $\Pi_C(i)$ and $\Pi_D(i)$ represent the average payoffs of a C and D player, respectively, in a population with i C players and $(N-i)$ D players. In the DG and the PGG, $\Pi_C(i) - \Pi_D(i)$ is always a negative constant, which is de-

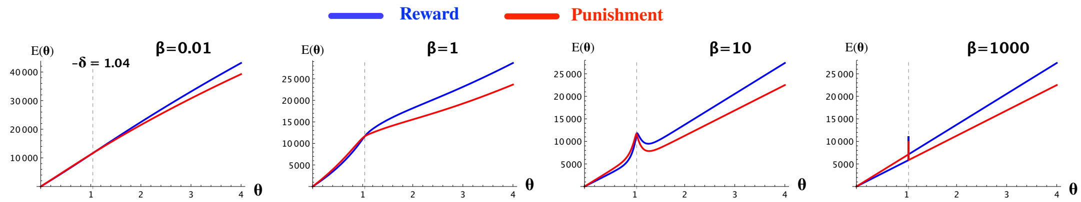

Figure 1: The expected total cost of investment E for reward and punishment, for varying θ and different values of β. Donation game: $b = 2$, $c = 1$; $N = 50$. When $\theta < -\delta$, punishment is more costly than reward, which is reversed when $\theta \geq -\delta$.

noted by $\delta < 0$. The entries n_{ij} of the so-called fundamental matrix $\mathcal{N} = (n_{ij})_{i,j=1}^{N-1} = (I - U)^{-1}$ of the absorbing Markov chain gives the expected number of times the population is in the state S_j if it is started in the transient state S_i (Kemeny and Snell, 1976). As a mutant can randomly occur either at S_0 or S_N, the expected number of visits at state S_i is: $\frac{1}{2}(n_{1i} + n_{N-1,i})$. The frequency of cooperation is given by $\frac{\rho_{D,C}}{\rho_{D,C}+\rho_{C,D}}$, where $\rho_{C,D}$ (resp. $\rho_{D,C}$) is the fixation probabilities of a C (respectively, D) player in a (homogeneous) population of D (respectively, C) players. Hence, this frequency of cooperation can be maximised by maximising

$$\max_{\theta} \left(\rho_{D,C}/\rho_{C,D}\right) = \max_{\theta} e^{\beta(N-1)(\delta+\theta)},$$

where the equality is obtained by simplifying the ratio $\rho_{D,C}/\rho_{C,D}$ following an established procedure (Nowak, 2006a). More generally, assuming that we desire to obtain at least an $\omega \in [0, 1]$ fraction of cooperation, i.e. $\frac{\rho_{D,C}}{\rho_{D,C}+\rho_{C,D}} \geq \omega$, then θ needs to satisfy the following lower bound (Han and Tran-Thanh, 2018)

$$\theta \geq \theta_0 = \frac{1}{(N-1)\beta} \log\left(\frac{\omega}{1-\omega}\right) - \delta.$$

Optimization problems. The expected total cost of interference for institutional reward and institutional punishment are respectively

$$E_r(\theta) = \frac{\theta}{2} \sum_{i=1}^{N-1} (n_{1i} + n_{N-1,i}) i,$$

$$E_p(\theta) = \frac{\theta}{2} \sum_{i=1}^{N-1} (n_{1i} + n_{N-1,i})(N-i).$$

In summary, we obtain the following cost-optimization problems of institutional incentives in stochastic finite populations: $\min_{\theta \geq \theta_0} E(\theta)$, where E is either E_r or E_p.

Main results. The main results of Duong and Han (2021) can be summarized as follows.

1. (infinite population limit)

$$\lim_{N\to+\infty} \frac{E(\theta)}{\frac{N^2\theta}{2}(\ln N + \gamma)} = \begin{cases} 1 + e^{-\beta|\theta - c|} & \text{for DG,} \\ 1 + e^{-\beta|\theta - c(1-\frac{r}{n})|} & \text{for PGG,} \end{cases}$$

where $\gamma = 0.5772...$ is the Euler–Mascheroni constant.

2. (weak selection limits) $\lim_{\beta\to 0} E(\theta) = N^2\theta H_N$, where $H_N = \sum_{i=1}^{N-1} \frac{1}{i}$ is the harmonic number.

3. (strong selection limit of E_r, E_p is similar)

$$\lim_{\beta\to+\infty} E_r(\theta) = \begin{cases} \frac{N^2}{2}\theta\left(\frac{1}{N-1} + H_N\right) & \text{for} \quad \theta < -\delta, \\ N^2\theta H_N \quad \text{for} \quad \theta = -\delta, \\ \frac{N^2}{2}\theta\left(1 + H_N\right) \quad \text{for} \quad \theta > -\delta. \end{cases}$$

4. There exists a threshold value β^* such that $\theta \mapsto E(\theta)$ is non-decreasing for all $\beta \leq \beta^*$ and is non-monotonic when $\beta > \beta^*$. As a consequence, for $\beta \leq \beta^*$

$$\min_{\theta \geq \theta_0} E(\theta) = E(\theta_0).$$

For $\beta > \beta^*$ and N is not too large ($N \leq N_0$ for some N_0), there exist $\theta_1 < \theta_2$ such that, $E(\theta)$ is increasing when $\theta < \theta_1$, decreasing when $\theta_1 < \theta < \theta_2$ and increasing when $\theta > \theta_2$. Thus, for $N \leq N_0$,

$$\min_{\theta \geq \theta_0} E(\theta) = \min\{E(\theta_0), E(\theta_2)\}.$$

5. $E_r(\theta) < E_p(\theta)$ for $\theta < -\delta, E_r(\theta) = E_p(\theta)$ for $\theta = -\delta$ and $E_r(\theta) > E_p(\theta)$ for $\theta > -\delta$.

Figure 1 demonstrates the behaviour of the cost function in different regimes of intensities of selection, when institutional reward is more or less costly than institutional punishment, as well as the phase transitions that occur when β is sufficiently large.

Summary and Outlook. We have summarized a recent theoretical analysis of the problem of optimizing cost of institutional incentives (for both reward and punishment) while guaranteeing a minimum amount of cooperation, in stochastic finite populations. In this context, institutional approaches have been widely adopted to study biological and artificial life systems (Andras et al., 2018; Jones et al., 2013; Smaldino and Lubell, 2014; Perret et al., 2019; Andras, 2020). This analysis provides new, fundamental insights into a cost-efficient design of institution-based solutions for promoting pro-social behaviours in social and artificial systems.

References

Andras, P. (2020). Composition of games as a model for the evolution of social institutions. In *Artificial Life Conference Proceedings*, pages 171–179. MIT Press.

Andras, P., Esterle, L., Guckert, M., Han, T. A., Lewis, P. R., Milanovic, K., Payne, T., Perret, C., Pitt, J., Powers, S. T., et al. (2018). Trusting intelligent machines: Deepening trust within socio-technical systems. *IEEE Technology and Society Magazine*, 37(4):76–83.

Chen, X., Sasaki, T., Brännström, Å., and Dieckmann, U. (2015). First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *Journal of The Royal Society Interface*, 12(102):20140935.

Cimpeanu, T., Han, T. A., and Santos, F. C. (2019). Exogenous rewards for promoting cooperation in scale-free networks. In *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, pages 316–323. MIT Press.

Cimpeanu, T., Perret, C., and Han, T. A. (2021). Cost-efficient interventions for promoting fairness in the ultimatum game. *Knowledge-Based Systems*, 233:107545.

Duong, M. H. and Han, T. A. (2021). Cost efficiency of institutional incentives for promoting cooperation in finite populations. *Proceedings of the Royal Society A*, 477(2254):20210568.

Han, T. A. (2013). *Intention Recognition, Commitments and Their Roles in the Evolution of Cooperation: From Artificial Intelligence Techniques to Evolutionary Game Theory Models*, volume 9. Springer SAPERE series.

Han, T. A., Pereira, L. M., and Lenaerts, T. (2015). Avoiding or Restricting Defectors in Public Goods Games? *J. Royal Soc Interface*, 12(103):20141203.

Han, T. A. and Tran-Thanh, L. (2018). Cost-effective external interference for promoting the evolution of cooperation. *Scientific reports*, 8(1):1–9.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

Jones, A. J., Artikis, A., and Pitt, J. (2013). The design of intelligent socio-technical systems. *Artificial Intelligence Review*, 39(1):5–20.

Kemeny, J. and Snell, J. (1976). *Finite Markov Chains*. Undergraduate Texts in Mathematics. Springer.

Maynard-Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.

Nowak, M. A. (2006a). *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, Cambridge, MA.

Nowak, M. A. (2006b). Five rules for the evolution of cooperation. *Science*, 314(5805):1560.

Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428:646–650.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge university press.

Perc, M., Jordan, J. J., Rand, D. G., Wang, Z., Boccaletti, S., and Szolnoki, A. (2017). Statistical physics of human cooperation. *Phys Rep*, 687:1–51.

Perret, C., Hart, E., and Powers, S. T. (2019). Being a leader or being the leader: The evolution of institutionalised hierarchy. In *Artificial Life Conference Proceedings*, pages 171–178. MIT Press.

Sasaki, T., Brännström, Å., Dieckmann, U., and Sigmund, K. (2012). The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences*, 109(4):1165–1169.

Sigmund, K. (2010). *The Calculus of Selfishness*. Princeton University Press.

Sigmund, K., Hauert, C., and Nowak, M. (2001). Reward and punishment. *P Natl Acad Sci USA*, 98(19):10757–10762.

Smaldino, P. E. and Lubell, M. (2014). Institutions and cooperation in an ecology of games. *Artificial life*, 20(2):207–221.

Traulsen, A., Nowak, M. A., and Pacheco, J. M. (2006). Stochastic dynamics of invasion and fixation. *Phys. Rev. E*, 74:11909.

Van Lange, P. A., Rockenbach, B., and Yamagishi, T. (2014). *Reward and punishment in social dilemmas*. Oxford University Press.

Wang, S., Chen, X., and Szolnoki, A. (2019). Exploring optimal institutional incentives for public cooperation. *Communications in Nonlinear Science and Numerical Simulation*, 79:104914.

Yang, C.-L., Zhang, B., Charness, G., Li, C., and Lien, J. W. (2018). Endogenous rewards promote cooperation. *Proceedings of the National Academy of Sciences*, 115(40):9968–9973.