

Forgiveness as a mean to resolve failures in cooperative agreements

Luis A. Martinez-Vaquero^{1,2,*}, The Anh Han³, Luis Moniz Pereira⁴ and Tom Lenaerts^{1,2}

¹AI lab, Computer Science Department, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

²MLG, Département d'Informatique, Université Libre de Bruxelles, Boulevard du Triomphe CP212, 1050 Brussels, Belgium

³School of Computing, Teesside University, Borough Road, Middlesbrough, UK, TS1 3BA

⁴NOVA Laboratory for Computer Science and Informatics, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

*l.martinez.vaquero@gmail.com

Making agreements on how to behave has been shown to be an evolutionarily viable strategy in one-shot social dilemmas (Han et al., 2013, 2015). Commitments are defined in this context as agreements to cooperate, with posterior compensations when any of the parties involved defects while the others honour the agreements. They offer an alternative pathway for the evolution of cooperation as opposed to punishment and rewards. Our analytical and numerical results (Martinez-Vaquero et al., 2015) reveal for the first time under which conditions revenge, apology, forgiveness and ostracism can evolve and deal with mistakes within ongoing agreements in the context of repeated games.

First we focus in iterated prisoner's dilemma, showing that, when the agreement fails, participants prefer to take revenge by defecting in the subsisting encounters. Then we introduce an apology-forgiveness mechanism in order to deal with mistakes. Figure 1 shows that when the compensation (γ) given upon apology is bigger than the cost of cooperating ($\gamma > 1$), proposers that cooperate during commitments, apologise when they defect by mistake and forgive when receiving an equivalent apology become the best strategists ($P, C, AllD, q = 1$). They reach a maximum continue dominating the population until becomes too high, leading to the situation where revenge, *i.e.* ($P, C, AllD, q = 0$), becomes once again the better choice. However, when the cost of apology is not high enough, fake proposers, *i.e.* ($P, D, AllD, q = 1$), take over. These fake proposers systematically exploit the apology-forgiveness mechanism, leading to the decrease of cooperation.

We have also studied commitments in repeated public good games where defective individuals can be expelled and ostracised at a given cost. We have shown that cooperators prefer not to ostracise individuals (or at least not for long time) unless the ostracising cost is very low and there is no compensation cost involved (in the form of fines for instance). Defectors, on the other hand, do not care about the reinsertion time. In a society dominated by defectors any ostracising time is equally likely to emerge since individuals do not collaborate to the public good and will eventually be expelled of the groups, making unsustainable any kind of

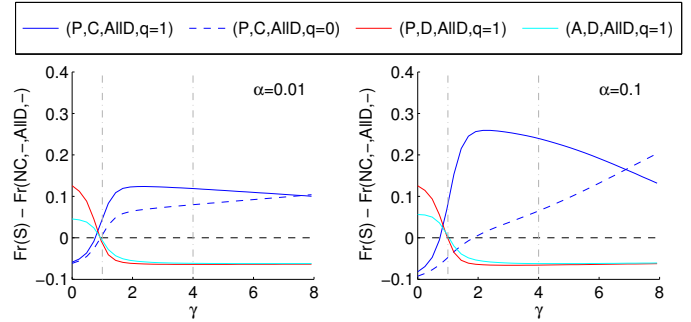


Figure 1: Stationary distribution of the main strategies with respect to the stationary distribution of the pure defectors as a function of the apology cost (γ) for different level of noise $\alpha = 0.01$ (left) and $\alpha = 0.1$ (right).

collaboration. Therefore the only way to maintain a cooperative society is if just after a defective action the individual pays a cost to repair her misconduct.

In short, forgiveness, in its different ways, is an evolutionarily viable strategy which plays an important role in inducing cooperation in repeated dilemmas both in iterated prisoner dilemma and in public good games. Apology needs to be sufficiently sincere, meaning not too low, not too high, in order for forgiveness to function properly.

References

- Han, T. A., Pereira, L. M., and Lenaerts, T. (2015). Avoiding or Restricting Defectors in Public Goods Games? *Journal of the Royal Society Interface*, 12(103):20141203.
- Han, T. A., Pereira, L. M., Santos, F. C., and Lenaerts, T. (2013). Good agreements make good friends. *Scientific Reports*, 3(2695).
- Martinez-Vaquero, L. A., Han, T. A., Pereira, L. M., Lenaerts, T., et al. (2015). Apology and forgiveness evolve to resolve failures in cooperative agreements. *Scientific reports*, 5.