Counterfactual Thinking in Cooperation Dynamics

Luís Moniz Pereira[1] and Francisco C. Santos[2]

[1]NOVA-LINCS and Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
[2]INESC-ID and Instituto Superior Técnico, Universidade de Lisboa, Portugal

Abstract. Counterfactual Thinking (CT) is a human cognitive ability studied in a wide variety of domains, namely Psychology, Causality, Justice, Morality, Political History, Literature, Philosophy, Logic, and AI. CT captures the process of reasoning about a past event that did not occur, namely what would have happened had this event occurred; or, otherwise, to reason about an event that did occur but what would ensue if it had not, or if another might have happened in its stead. An example: *Lightning hits a forest and a devastating forest fire breaks out. The forest was dry after a long hot summer and many acres were destroyed*. A counterfactual thought is: *If only there had not been lightning, then the forest fire would not have occurred*. Given the wide cognitive empowerment of CT in the human individual, the question arises of how the presence of individuals with CT-enabled strategies affects the evolution of cooperation in a population comprising individuals with diverse strategies. The natural locus to examine this issue is Evolutionary Game Theory (EGT). In EGT, a simple CT can be exercised after knowing one's resulting payoff following a single playing step with a co-player. It employs the counterfactual: *Had I played differently, would I have obtained a better payoff than I did?* This information can be easily obtained by consulting the game's payoff matrix, assuming the co-player would make the same play, that is, other things being equal. In the positive case, the CT player will next adopt the alternative play strategy. Here we compare the population dynamics emerging from the interplay between CT and Social Learning (SL) when individuals face cooperation dilemmas. Preliminary results suggest that CT can foster coordination in collective action problems. Our initial approach begs further developments: consideration of direct reciprocity; allowing strategy revision between CT and SL; using CT over a history of plays or a corpus of games; comparing CT to other reasoning methods, like forward reasoning, best response, intention recognition; susceptibility to noise; etc.

Acknowledgements. We are grateful to The Anh Han and Tom Lenaerts for comments. This work was supported by FCT-Portugal/MEC through grants NOVA-LINCS UID/CEC/04516/2013, INESC-ID UID/CEC/50021/2013, PTDC/EEI-SII/5081/2014, and PTDC/MAT/STA/3358/2014.

Counterfactual Thinking (CT) is a human cognitive ability studied in a wide variety of domains, namely Psychology, Causality, Justice, Morality, Political History, Literature, Philosophy, Logic, and AI. For surveys see [1-7]. Given the wide cognitive empowerment of CT in the human individual, the question arises of how the presence of individuals with CT-enabled strategies affects the evolution of cooperation in a population comprising individuals of diverse interaction strategies. The natural locus to examine this issue is Evolutionary Game Theory (EGT) [8], given the amount of extant knowledge concerning different types of games, strategies, and techniques for the evolutionary characterization of such populations of individual players. In the context of EGT, individuals revise their strategy looking at the success and actions of others. Yet, contrary to social learning, an agent may instead imagine how an outcome could have turned out, if she would have decided differently, and revise strategy. Here we propose a simple model to study the impact on cooperation of having a fraction of agents resorting to such counterfactual reasoning in a population of social learners.

In EGT, a simple CT can be exercised after knowing one's resulting payoff following a single playing step with a co-player. It employs the counterfactual thought: *Had I played differently, would I have obtained a better payoff than I did?* This information can be easily obtained by consulting the game's payoff matrix, assuming the co-player would have made the same play, that is, other things being equal. In the positive case, the CT player will next adopt the alternative play strategy.

A more sophisticated CT would search for a counterfactual play that improves not just one's payoff, but one that also contemplates the co-player not being worse off, for fear the co-player will react negatively to one's change of strategy. More sophisticated still, the new alternative strategy may be searched for taking into account that the co-player also possesses CT ability. And the co-player might too employ a Theory of Mind (ToM)-like CT up to some level.

In this extended abstract we consider only the first model, for simple (egotistic) CT. CT can be envisaged as a form of strategy update, akin to debugging and best-response rule, in the sense that *if my actual play move was not conducive to a good payoff, then, after having known the co-player's move, I can imagine how I would have done better had I made a different strategy choice*. In EGT, a frequent standard form of learning is so-called Social Learning (SL). It basically consists in switching one's strategy by imitating the strategy of a more successful individual in the population, compared to ours.

Let us consider a population of Z agents interacting in groups of N individuals, who can be either Cooperators (C) or Defectors (D); the k Cs in N contribute a cost c to the public good, whereas Ds refuse to do so. The accumulated contribution is multiplied by an enhancement factor F, and the result equally distributed among all individuals of the group, irrespectively whether they contributed or not. Moreover, we often find situations where a minimum number of Cs are needed within a group to create any sort of collective benefit [9, 10]. This can be modeled through the addition of a coordination threshold M [9], leading to the following payoffs of Cs and Ds: $P_D(k) = H(k-M)kFc/N$ and $P_C(k) = P_D(k) - c$, where $H(x)=1$ if $x \geq 0$, and $H(x)=0$ otherwise. This game is commonly referred as N-person Stag-Hunt game [9]. As in standard EGT, we consider well-mixed populations where all agents are equally likely to interact with each other. Thus, the success (or fitness) of an agent results from the average payoff obtained from randomly sampling groups of size N (see [9, 11] for details).

Let us consider finite well-mixed populations of Z interacting agents, of either SL or CT agents only, that revise their behaviors accordingly, both implemented by means of a stochastic update rule. If she resorts to social learning, i will imitate a randomly chosen individual j, with a probability p that augments with the increase in fitness difference between j and i, given by f_i and f_j respectively. This probability may be conveniently written in terms of the Fermi distribution $p_{SL} = \left[1 + e^{-\beta_{SL}[f_j - f_i]}\right]^{-1}$, in which β_{SL} translates here into noise associated with errors in the imitation process [12]. Hence, successful individuals might be imitated and the associated strategy will spread in the population. Given the above assumptions, considering a 2-strategy model (strategies C and D), it is easy to write down the probability to change the number k of Cs (by ± one at each time step) in a population of Z-k Ds in the context of social learning: $T_{SL}^{\pm}(k) = \frac{k}{Z}\frac{Z-k}{Z}\left[1 + e^{\mp\beta_{SL}[f_B(k) - f_A(k)]}\right]^{-1}$. Differently, agents that resort to CT assess alternatives to their present returns, had they used the alternative, *contrary* to what actually took place. Agents imagine how the outcome could have worked if her decision (or strategy, in this case) would have been different. In its simplest form, this can be modeled as an incipient form of myopic best response rule or logit rule [13, 14]. In the case of CT, an individual i adopting a strategy A will switch to B with a probability $p_{CT} = \left[1 + e^{-\beta_{CT}[f_i^B - f_i^A]}\right]^{-1}$ that increases with the fitness difference between the fitness she would have had if she had played B (f_i^B), and the fitness she actually got by playing A (f_i^A). Once again, β_{CT} translates here into noise associated with guessing the fitness values. As before, one may write down the probability to change the number k of Cs by plus or minus one as $T_{CT}^{+}(k)=\frac{Z-k}{Z}\left[1+e^{-\beta_{CT}\left[f_i^C-f_i^D\right]}\right]^{-1}$ and $T_{CT}^{-}(k)=\frac{k}{Z}\left[1+e^{-\beta_{CT}\left[f_i^D-f_i^C\right]}\right]^{-1}$. The above transition probabilities can be used to assess the most probable direction of evolution in SL and CT. This is given by a learning gradient (often called gradient of selection [9, 11] in the case of SL), $G_{SL}(k)=T_{SL}^{+}(k)-T_{SL}^{-}(k)$ and $G_{CT}(k)=T_{CT}^{+}(k)-T_{CT}^{-}(k)$, respectively. When $G_{SL}(k)>0$ and $G_{CT}(k)>0$ ($G_{SL}(k)<0$ and $G_{CT}(k)<0$), time evolution is likely to act to increase (decrease) the number of Cs. We further assume that with probability μ individuals may switch to a randomly chosen strategy, freely exploring the space of possible behaviors.

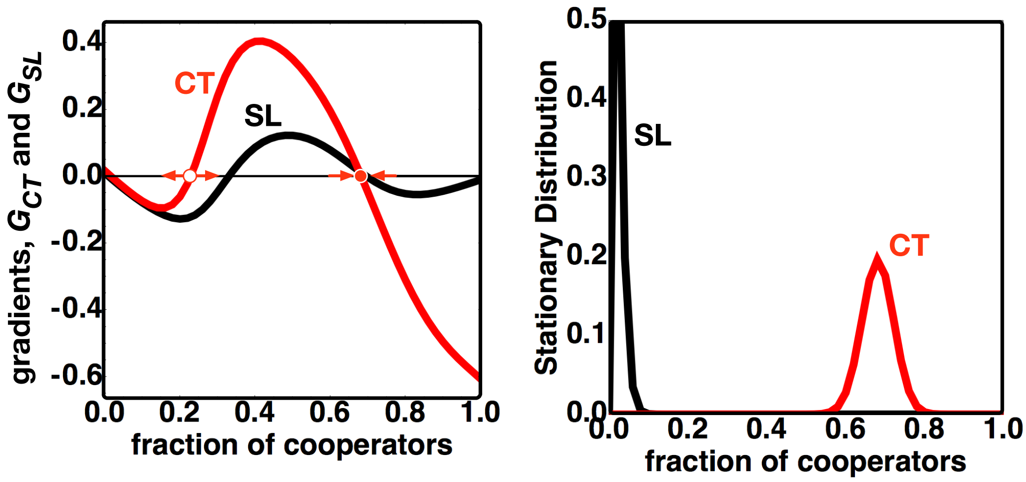

Fig. 1. Left panel: Learning gradients ($G_{SL}(k/Z)$, black line, and $G_{CT}(k/Z)$, red line) for the N-person SH game [9] (Z=50, N=6, F=5.5. M=3, c=1.0, μ=0.01, $\beta_{SL}=\beta_{CT}$=5.0). Empty and full circles represent the finite population analogue of unstable and stable fixed points, respectively, in infinite populations ($Z\rightarrow\infty$). Right panel: Stationary distribution of the Markov processes created by the transition probabilities pictured in the left panel; it characterizes the prevalence in time of each fraction of cooperators in finite populations. (For details on the methodology see, e.g., [11]).

In Figure 1a, we illustrate how CT allows for the creation of new playing strategies, absent before in the population, since new strategies can appear spontaneously based on individual reasoning. By doing so, CT interestingly leads to different results if compared to SL. In this particular scenario, it is evident how CT may facilitate coordination of action, as individuals can reason on the sub-optimal outcome associated with non-reaching the coordination threshold, and individually react to that. In Figure 1b, we show the stationary distribution of the Markov chain associated with the transition probabilities indicated above, showing how cooperation can benefit from CT. The stationary distribution characterizes the prevalence in time of each fraction of cooperators (k/Z). In this particular configuration, it is shown how in SL (black line), the population spends most of the time in low values for the fraction of cooperators. Whenever CT is allowed, cooperation is maintained most of the time.

In this very simple model, we disallow SLs to learn from CTs, and vice-versa, since the two populations are not mixed. In our hybrid population analytic model, not shown, we allow for that. The cases where CTs have access to a lengthier own record of plays, rather than just the last one, creates a time-dependence that should be addressed in the future resorting to numerical computer simulations.

References

[1] Mandel D R, Hilton D J and Catellani P 2007 *The psychology of counterfactual thinking* (Routledge)

[2] Birke D, Butter M and Köppe T 2012 *Counterfactual Thinking-Counterfactual Writing* vol 12 (Walter de Gruyter)

[3] Roese N J and Olson J M, eds 2014 *What might have been: The social psychology of counterfactual thinking* (Psychology Press)

[4] Wohl V 2014 *Probabilities, Hypotheticals, and Counterfactuals in Ancient Greek Thought* (Cambridge University Press)

[5] Dietz E-A, Hölldobler S and Pereira L M 2015 On Conditionals. In: *Global Conference on Artificial Intelligence (GCAI 2015),* ed G Gottlob, *et al.* (Tbilisi, Georgia: Citeseer) pp 79-92

[6] Nickerson R 2015 *Conditional Reasoning: The Unruly Syntactics, Semantics, Thematics, and Pragmatics of "if"* (Oxford University Press)

[7] Pereira L M and Saptawijaya A 2017 Counterfactuals, Logic Programming and Agent Morality. In: *Applications of Formal Philosophy*, eds. R Urbaniak and G Payette (Springer) pp 25-53

[8] Sigmund K 2010 *The Calculus of Selfishness* (Princeton Univ. Press)

[9] Pacheco J M, Santos F C, Souza M O and Skyrms B 2009 Evolutionary dynamics of collective action in N-person stag hunt dilemmas *Proc R Soc Lond B* **276** 315-21

[10] Souza M O, Pacheco J M and Santos F C 2009 Evolution of cooperation under N-person snowdrift games *J Theor Biol* **260** 581-8

[11] Santos F C and Pacheco J M 2011 Risk of collective failure provides an escape from the tragedy of the commons *Proc Natl Acad Sci USA* **108** 10421-5

[12] Traulsen A, Nowak M A and Pacheco J M 2006 Stochastic dynamics of invasion and fixation *Phys. Rev. E* **74** 011909

[13] Sysi-Aho M, Saramäki J, Kertész J and Kaski K 2005 Spatial snowdrift game with myopic agents *The European Physical Journal B-Condensed Matter and Complex Systems* **44** 129-35

[14] Perc M, et al. 2017 Statistical physics of human cooperation *Physics Reports* **687** 1-51