Chapter 24: The Messianic Effect of Pathological Altruism

Jorge M. Pacheco[1,3,4] and Francisco C. Santos[2,3,4]

[1] Departamento de Matemática e Aplicações, Universidade do Minho, 4710 - 057 Braga, Portugal
[2] CENTRIA, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.
[3] ATP-group, CMAF, Complexo Interdisciplinar, P-1649-003 Lisboa Codex, Portugal,
[4] GADGET, Apartado 1329, 1009-001, Lisboa, Portugal.

Key Concepts

- Without additional mechanisms, cooperation is not an evolutionarily viable behavior as the *tragedy of the commons* often emerges as the final doomsday scenario.
- In a black and white world in which individuals' actions are limited to cooperate or to defect, pathological altruists can be seen as obstinate cooperators, who go at all lengths to maintain their behavior.
- Pathological altruists cooperate indiscriminately, being unmoved by the temptations of greed and fear that leads to defection.
- A single pathological altruist can obliterate the evolutionary advantage of defectors, letting others ignore the temptation to cheat and become, themselves, cooperators. Hence, they generate a messianic effect, which spreads through the entire community.
- Pathological altruists catalyze social cohesion, as their presence benefits the entire community even when defection remains as the single rational option and individuals act in their own selfish interest.

Introduction

Humans live in large societies characterized by exchange and cooperation between individuals who, in the majority of cases, are not kin-related. Close examination reveals that humans actually cooperate more often than would be expected from evolutionary game theory as modeled in terms of the classic Prisoner's Dilemma (Axelrod & Hamilton, 1981; Hofbauer & Sigmund, 1998; Maynard-Smith, 1982). Prisoner's Dilemma is rooted on the assumption that the act of cooperation entails a certain cost c, which need not be a monetary cost. The recipient of a cooperative act receives a benefit b. Quantitatively, the magic of cooperation relies on the fact that $b > c$.

In a black and white world in which people can only behave as cooperators or defectors, one of only four possible outcomes takes place when two individuals interact. When both *cooperate* (C), each receives a benefit b but also experiences a cost c; hence each receives a net profit of b-c. When both *defect* (D), neither player receives cost nor benefit. Lastly, if one player cooperates while the other defects, then D receives a benefit without a cost, whereas C experiences a cost with no benefit. These four entries fill in what is known in game theory as the *payoff matrix*:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array}$$

with $R=(b-c)$, $T=b$, $S=-c$ and $P=0$. These four entries satisfy the ranking order $T>R>P>S$, which is the hallmark of a Prisoner's Dilemma game. The fact that mutual cooperation is always better than mutual defection implies that $R>P$. When $T>R$, one may think of *greed*, as an individual is tempted to play *D* towards a *C* (Macy & Flache, 2002). Indeed, in the absence of *greed* ($T<R$), the dilemma is relaxed from a pure defector dominance game into a coordination game, termed the *Stag-Hunt Dilemma* (Skyrms, 2004). In this case, only the *fear* of being cheated on by a defector ($P>S$) provides a reason for defecting instead of cooperating (Macy & Flache, 2002). But there is yet another scenario—that in which *fear* is removed from the Prisoner's Dilemma so that *greed* becomes the only reason to defect. This dilemma then becomes a coexistence game, known as the *Chicken, Hawk-Dove* or *Snowdrift Dilemma* (Maynard-Smith, 1982).

In sum, then, Prisoner's Dilemma emerges as the most stringent of the social dilemmas captured in terms of symmetric, one-shot, two-player games. It is the stringent Prisoner's Dilemma, in its cost-benefit version (the parameterization above), that constitutes the hallmark of most studies carried out to date addressing the evolution of cooperation (Nowak, 2006a, 2006b; Taylor, Day, & Wild, 2007).

A mathematical model of pathological altruism

In keeping with such studies, we shall also adopt the Prisoner's Dilemma, and consider a finite population, small enough to make it equally likely that anyone in the population could interact with anyone else (Dunbar, 2003). This is the commonly encountered *well-mixed assumption* (known as the *mean field approximation* in physics) (Hofbauer & Sigmund, 1998). Under such circumstances, cooperators are always at a disadvantage when compared with defectors, and natural selection favors the increase of *D*s at the expense of *C*s. This is related to the fact that the payoff for *C*s (interpreted as fitness or social success in evolutionary game theory) is lower than that of *D*s. For a population of size *N* with *k* *C*s, the average payoff of *C*s and *D*s is

$$\Pi_C(k) = \frac{k}{N}R + \frac{N-k}{N}S = \frac{k}{N}b - c \qquad (1)$$

$$\Pi_D(k) = \frac{k}{N}T + \frac{N-k}{N}P = \frac{k}{N}b \qquad (2)$$

(for $R=(b-c)$, $T=b$, $S=-c$ and $P=0$, ignoring residual self-interaction corrections), and, since $T>R>P>S$ we immediately see that *C*s do worse than *D*s independently of *k*, which means *D*s ultimately dominate unconditionally the evolutionary dynamics in Prisoner's Dilemma (see Figure 1d).

Besides individual fitness, evolutionary dynamics relies on a process by which individuals revise their strategic behavior. Here we adopt a popular stochastic update known as the *pairwise comparison rule* (Szabó & Tőke, 1998; Traulsen, Nowak, & Pacheco, 2006): At each time step an individual *i* will adopt the strategy of a randomly chosen individual in the population *j* with a probability that increases with the increase in payoff difference between *j* and *i*. Hence, successful behaviors will be imitated and spread in the population.

This probability is conveniently written in terms of the so-called Fermi distribution (from statistical physics) $F\left[\Pi_j(k) - \Pi_i(k)\right] = \left[1 + e^{-\beta[\Pi_j(k) - \Pi_i(k)]}\right]^{-1}$, in which $\Pi_i(k)$ and $\Pi_j(k)$ are the payoffs of individuals *i* and *j*, respectively defined in equations (1) and (2), and β (an inverse temperature in physics) translates here into noise associated with errors in decision making.

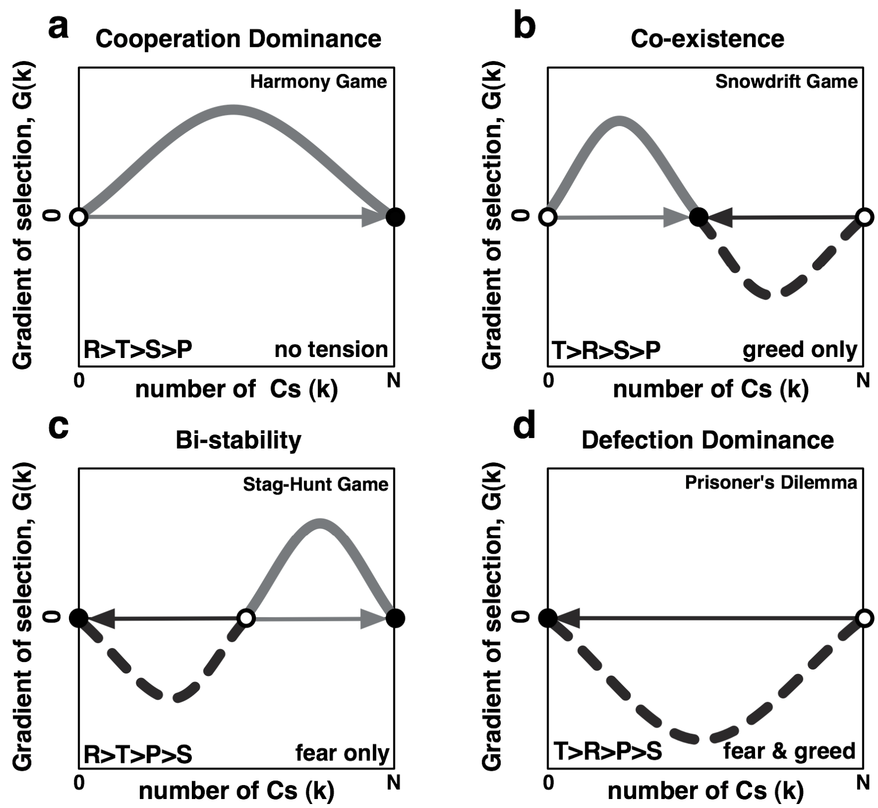


Figure 1. The behavioral dynamics of a population can be studied by analyzing the sign of the gradient of selection $G(k)$, defined as the difference between the probability of increasing in the number of cooperators, $T^+(k)$ versus that of decreasing, $T(k)$ for each value k of cooperators in a population of size N. Whenever $G(k)>0$, cooperators will have an advantage over defectors, increasing their fraction in the population. On the contrary, if $G(k)<0$, evolution will promote the increase of defectors. Here we depict the dynamics of symmetric two-person one-shot dilemmas using gradients of selection. When cooperators have an advantage irrespectively of k, a Harmony Game is obtained (**panel a**). In this situation, individual and collective interests always coincide, and hence there is no dilemma. In the remaining three cases, individual and collective interests no longer coincide: *Co-existence* can be promoted (**panel b**) in situations in which a minority of individuals adopting a given behavior gain an advantage, losing this advantage when they become abundant. Hence, there is an *internal equilibrium* which is stable, represented in **panel b** by a solid circle. A *coordination dilemma* emerges whenever the opposite occurs (**panel c**) — the *internal equilibrium* becomes *unstable* (represented by an open circle in **panel c**). Finally, whenever cooperation is always a disadvantage we obtain a Prisoner's Dilemma situation in which cooperators have no chance to survive (**panel d**). The figure also illustrates how *greed* and *fear* operate in the various dilemmas. *Greed* alone (temptation to defect, $T>R$) results in a *stable equilibrium* (**panel b**). *Fear* alone (fear of being cheated upon, $P>S$) results in an *unstable equilibrium* (**panel c**). *Greed* and *fear* are both present as we move to **panel d**, with its combination of *stable* and *unstable equilibrium points*.

For high values of β we obtain pure imitation dynamics commonly used in cultural evolution studies, whereas for $\beta \rightarrow 0$, selection becomes so weak that evolution proceeds by random drift.[1] Under such a stochastic dynamics, one can compute the probabilities $T^+(k)$ and $T^-(k)$ for the number of *C*s in the population to grow or diminish by a single cooperator in a given time step. Assuming there are *k* *C*s in a population of size *N*, we may write

[1] Decreasing values of β may be thought of as increasing the likelihood that someone who actually wants to help fails to do so. For instance, someone comes across a beggar and wants to give him some money but realizes that has forgotten the purse. In other words, an example of a cooperator who fails to act accordingly. β is also related, indirectly, to the issue of bounded rationality — sometimes, just by chance, one does not do what one is supposed to do rationally. Moreover, β measures errors in the imitation process related with the fact that often individuals face difficulties in assessing the success (or not) of others. This may lead individuals to change their behavior to something which is, in fact, worse than their previous choice.

$$T^{+}(k)=\frac{N-k}{N}\frac{k}{N}F\left[\Pi_{C}(k)-\Pi_{D}(k)\right] \qquad (3)$$

$$T^{-}(k)=\frac{k}{N}\frac{N-k}{N}F\left[\Pi_{D}(k)-\Pi_{C}(k)\right] \qquad (4)$$

such that the sign of the gradient of selection $G(k)=T^{+}(k)-T^{-}(k)$ indicates whether evolution favors the increase ($G(k)>0$) or decrease ($G(k)<0$) of *C*s in the population. In Figure 1 we show the typical profile of $G(k)$ for the Prisoner's Dilemma and other social dilemmas. Given the stochastic nature of the dynamics introduced, combined with the finite size of the population, the end states of evolution are inevitably monomorphic, that is, populations will be entirely comprised of Cooperators only or Defectors only, which become absorbing states of the evolutionary dynamics. Only in infinite populations can polymorphic states become stable. Yet, as shown in Figure 1, even in finite populations natural selection may lead populations to spend most of their time in polymorphic states, associated with the internal roots of $G(k)$. Hence we employ, for finite populations, the same nomenclature which is strictly correct only in infinite deterministic dynamics, using an italic font to emphasize this association. With this proviso in mind, our discussion should cause no confusion.

What happens if we now introduce a small number of pathological altruists (*PA*) in this population? Unlike conventional *C*s, *PA*s do not imitate or let themselves be influenced by anyone – they are *obstinate C*s. Hence, and similar to *C*s, they suffer the exploitation of *D*s while benefiting from the cooperation of *C*s (and other *PA*s, if present). Although they do not imitate anyone, their altruistic behavior can be imitated by others – those who do so will be *D*s who become *C*s, given that, from the outset, *PA*s and *C*s are indistinguishable. Let $p \leq N$ be the (fixed) number of *PA*s in the population. If $k=k'+p$, where k' is the number of "conventional" *C*s in the population, then the payoff of *C*s and *D*s is still given by equations (1) and (2), whereas the transition probabilities now read

$$T_{PA}^{+}(k)=\frac{N-k}{N}\frac{k}{N}F\left[\Pi_{C}(k)-\Pi_{D}(k)\right] \qquad (5)$$

$$T_{PA}^{-}(k)=\frac{k-p}{N}\frac{N-k}{N}F\left[\Pi_{D}(k)-\Pi_{C}(k)\right] \qquad (6)$$

where $0 \leq p \leq k \leq N$.

Evolutionary dynamics of pathological altruists

Comparison of equations (3) and (4) with equations (5) and (6) shows a subtle difference rooted in the profound changes introduced by the existence of *PA*s in the population—no matter how few *PA*s are introduced. The pre-factors of the Fermi function no longer coincide in equations (5) and (6). Instead, the symmetry is broken by the appearance of an additional term in *p* due to the presence of *PA*s. As we show below, this term is capable of disrupting the unconditional dominance of *D*s portrayed in Figure 1d for the Prisoner's Dilemma. Indeed, this additional factor provides an overall net positive contribution to $G(k)$, with important consequences in the overall evolutionary dynamics of the population. Figure 2 provides a concrete example of the impact of *PA*s in a population of N=200 individuals. In particular, the fact that $G(p)=\frac{p}{N}\frac{1-p}{N}\frac{1}{1+e^{-\beta c}}>0$ (for all p>0) means that a single *PA* is sufficient to reverse the direction of natural selection, compared to the conventional Prisoner's Dilemma, to instead favor an increase in the number of cooperators when these are rare.

In essence, then, the presence of pathological altruists means that *C*s no longer tend to go extinct, as in the standard Prisoner's Dilemma (*p*=0 in Figure 2a). Instead, natural selection now drives the population into an internal *polymorphic equilibrium* characterized by the *coexistence* of *C*s and *D*s in the population. As Figure 1 revealed, such a *coexistence equilibrium* was possible due to *greed* alone, but the presence of *PAs* now renders *fear* and *greed* no longer sufficient to stop cooperators from surviving in a population in which at least one *PA* appears. In fact, for given values of *c*, β and *p*, *equilibrium* is attained at

$$k^* = \frac{p}{1-e^{-\beta c}}. \qquad (7)$$

This is a remarkable result. The presence of *p PA*s induces an internal *stable equilibrium* in the evolutionary dynamics. More importantly, this *equilibrium* occurs for a value *k**>*p*, as shown in Figure 2b, a result which does not depend sensitively on the specific value of the *cost-to-benefit* ratio of cooperation. In other words, the presence of *PA*s catalyzes the appearance of standard cooperators in the population. It is also noteworthy that all this happens despite the fact that *C*s and *PA*s have a lower fitness than *D*s.

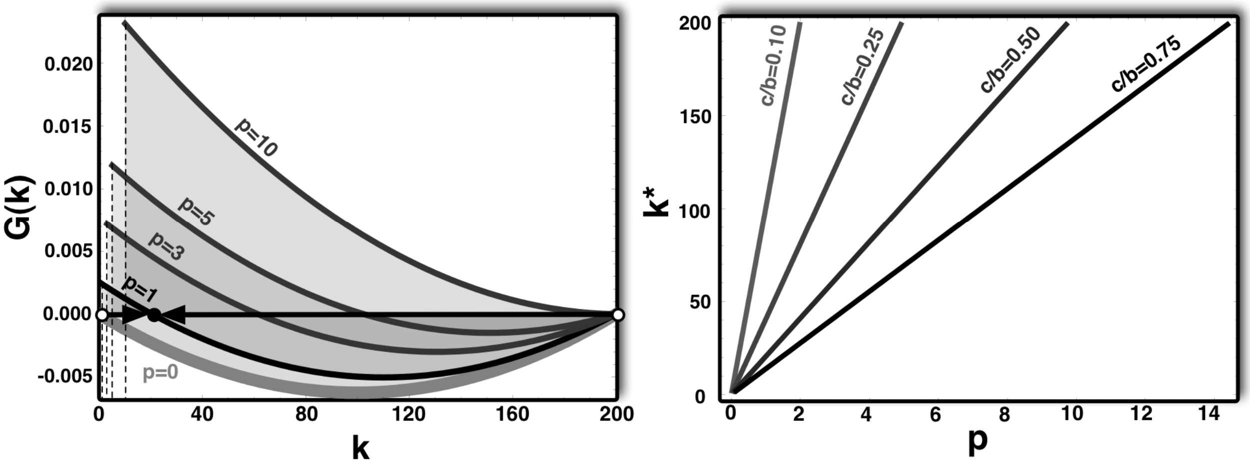


Figure 2. Left panel. Dynamics of cooperation under a Prisoner's Dilemma (c/b=0.5) for different numbers of *PA*s in a population of *N*=200 individuals. A single *PA* (*p=1*) is able to transform the original Prisoner's Dilemma into a *co-existence* game with a *stable equilibrium* in the interval $p < k \leq N$. In this particular case, for *p*=1 we obtain (*β=0.1*) *k*=14* (see equation 7 and right panel), whereas for *p*=3 we obtain *k*=61*. These results should be compared with the conventional dynamics corresponding to *p*=0, in which case *C*s are not evolutionary viable. The vertical dashed lines indicate the minimum value *k=p*. **Right panel.** *Stable equilibria k** for the same conditions as in the left panel and different values of *c/b* and *p*. A small number of *PA*s is able to create a spectacular boost of cooperators in the population, providing evidence of the messianic effect of *PA*s, a process that occurs independently of the specific value of the cost-to-benefit ratio associated with the act of cooperation.

From a mathematical perspective, and to the best of our knowledge, this is the first time one obtains an evolutionary dynamics in a well-mixed population in which the *internal equilibria do not* coincide with the zeroes of $\Pi_D(k) - \Pi_C(k)$. For general symmetric two-person games, this difference depends on the number of *k* cooperators (pathological or not) and it is the possibility that this difference becomes zero that leads to the appearance of *internal equilibria*, *stable* or *not*. However, in the present case, and for the particular (so-called "benefit-cost") parameterization of the Prisoner's Dilemma adopted, $\Pi_D(k)-\Pi_C(k)=c$ for all *k*, and hence the evolutionary viability of *C*s in the presence of *PA*s is due to the modified nature of the evolutionary dynamics, which no longer follows a standard replicator-like equation. This is easily understood when we take the (unrealistic) limit of infinite, well-mixed populations. To this end, we define, in the usual sense, $x=k/N$ as the fraction of *C*s (and *PA*s), and $\phi = p/N \le x$ as the fraction of *PA*s in the population, such that the corresponding fraction of *D*s becomes $1-x$. Taking the limit $N \to \infty$ and maintaining both *x* and ϕ constant leads to the following differential equation (Traulsen, et al., 2006)

$$\dot{x} = x(1-x)\tanh\left[\beta\!\big/_2\left(\Pi_C(x)-\Pi_D(x)\right)\right] + \phi(1-x)F\left[\Pi_D(x)-\Pi_C(x)\right].$$

The first term on the right-hand side is nothing but the standard modified replicator dynamics equation resulting from the pairwise comparison rule (Traulsen, et al., 2006), adopted for strategy update, and governed by the fitness difference between *C*s (and *PA*s), and *D*s. The second term results from the presence of *PA*s in the population, and is due to the inability of the evolutionary dynamics to reach values of *x* satisfying $x \le \phi$. More important, however, is the fact that $\dot{x}(\phi) > 0$, transforming $x=\phi$ into an *unstable fixed point*, promoting the appearance of *C*s in the population.[2]

Discussion

The present model studies the impact of a fixed amount of *PA*s on the evolutionary dynamics of a finite, well-mixed population. *PA*s are obstinate cooperators who maintain their strategies irrespective of any stimuli to change that may surround them. We find that the presence of *PA*s in a population of size *N* leads the population to spend most of the time in a polymorphic composition in which the *equilibrium* number of *C*s is given by

$$k_C^* = k^* - p = \frac{p}{e^{\beta c}-1}.$$

Hence,
i) the more *PA*s in the population,
ii) the weaker the force of natural selection or
iii) the smaller the cost of cooperation,

the larger the incidence of cooperators in the population. In fact, whenever the product βc satisfies $\beta c < -\ln(1-p/N)$, natural selection will favor the extinction of defectors. This is a remarkable effect in what concerns the impact of *PA*s in the evolutionary dynamics of the population. What is the intuition behind this result?

As becomes clear from the discussion above, the fitness of an individual results from her interaction with her peers. These interactions clearly favor *D*s, as individuals engage here in a Prisoner's Dilemma. However, the evolutionary dynamics of the strategies within the population depend only partially on individual fitness. Indeed, for obstinate *PA*s what difference does the fitness of others make

---

[2] Metaphorically, this might explain why dictators go to any lengths to purge and eliminate those who speak out against them. But in those cases, so many factors contribute to such perversion that it is difficult to disentangle the effects of pathological altruists. In modern times, attempts to control the media by governments and the existence of those who resist such attempts parallels, to some extent, the classic dictator example.

if the PAs themselves will never deviate from their altruistic behavior? An easy means to disentangle the roles played by fitness and *strategy update* is to view the fate of individuals as proceeding along the links of one or more complex networks.

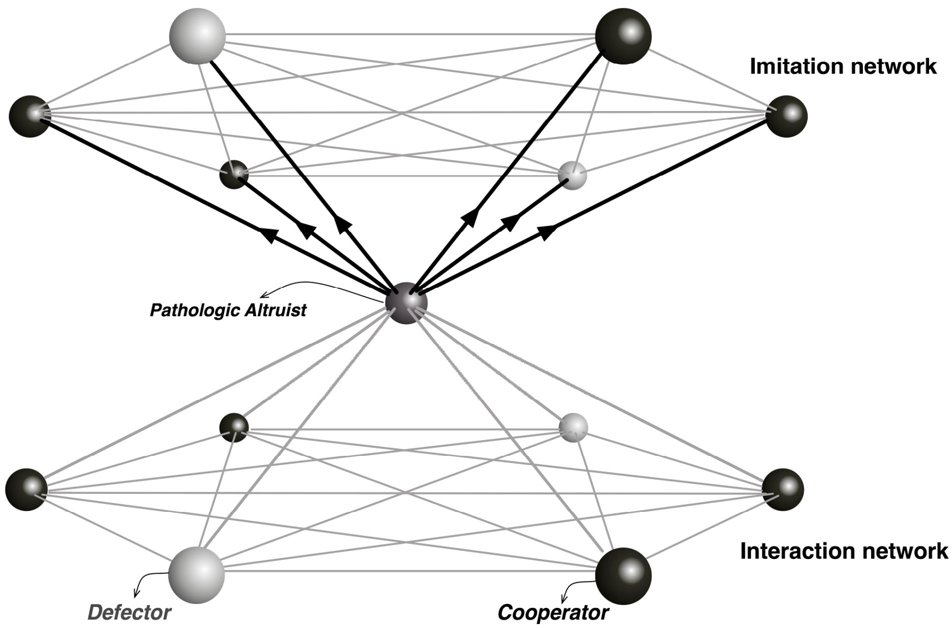


Figure 3. The figure illustrates a well-mixed population of seven individuals, in this case four *C*s (black) two *D*s (grey) and one *PA* (dark grey). The evolutionary dynamics is the result of *i)* interactions between individuals which proceed along the links of the interaction network (bottom, a complete graph of bi-directional links) and *ii)* behavior update which proceeds along the links of the imitation network (top, a complete graph of bi-directional links except those that emerge from the single *PA*. Here we adopt the notation that bidirectional links have no arrows, in contrast with directional links. Because links to *PA*s in the imitation network are *not* bi-directional, reflecting the obstinacy of *PA*s who never change behavior, the evolutionary dynamics of a population in the presence of *PA*s is profoundly affected by them: Their influence induces the emergence of *C*s. As a visualization aid, nodes of the networks depicted have different sizes (bigger are meant to be closer) merely to induce a rudimentary sense of perspective to the picture.

Under the well-mixed assumption, everyone is connected to everybody else and will freely interact with everybody. Hence, we can define an interaction network, illustrated at the bottom of Figure 3, associated with a bi-directional complete graph, where individuals occupy the nodes of the graph, and the links between nodes define who interacts with whom. On the other hand, and inspired by the work described in (Ohtsuki, Pacheco, & Nowak, 2007) and (Ohtsuki, Nowak, & Pacheco, 2007), we can also define a second graph, the so-called reproduction, update or imitation graph, represented on the top of Figure 3. Similar to the interaction network, individuals occupy nodes (the same nodes, as the individuals are the same in both graphs) but now the links are no longer bidirectional, as in this case some individual may use another as a role model without the reverse being true. This is precisely the case of *PA*s, who may be role models of all non-*PA*s in the population, but accept no role models themselves. That is, *PA*s are effectively disconnected in the imitation network, although they remain fully connected with everybody else through the interaction graph.

Because of this peculiar topology of the imitation graph, the presence of *PA*s induces a symmetry breaking which is ultimately responsible for their *messianic effect* in the population as a whole, paving the way for cooperators to thrive. Depending on the value of the cost implicit in each act of cooperation, as well as on how strong natural selection leads individuals to change their strategy, the presence of rare (e.g., a single individual, see Figure 2) *PA*s may be enough to change the evolutionary dynamics from one in which *C*s become extinct into another in which *C*s dominate. This is a remarkable effect of *PA*s that can be rationalized in terms of their strong role in the imitation sector of the evolutionary dynamics. This is more so whenever selection is weak.

As argued elsewhere (Nowak, 2006a) the Prisoner's Dilemma game considered here is but one of the many evolutionary game theory games that "individuals" engage in. That is to say, in both game theory and real life, individuals have many interaction networks. And in these different networks, and even in the same network at different times, individuals play different games, some of which involve cooperation, some not. All of these interactions ultimately contribute to the fitness value of each individual. As such, when one concentrates on a single game, as we did here, it is natural to assume that fitness changes resulting solely from this game will be small or, equivalently, selection pressure due to this game alone will be weak. But this means, then, that random drift will dominate (or other games will be perhaps more important, which is not of interest here), and the weaker the effect of the game on fitness, the stronger the role of obstinate *PA*s. Hence, one expects *PA*s to introduce profound changes in the evolutionary dynamics of well-mixed communities.

In view of the discussion so far, the question remains regarding the origin of *PA*s and how they may actually emerge in a population. As we (Santos, Santos, & Pacheco, 2008; Van Segbroeck, Santos, Lenaerts, & Pacheco, 2009) and others (McNamara, Barta, Fromhage, & Houston, 2008; McNamara, Barta, & Houston, 2004) have argued at length, humans are prone to explore new forms of behavior, and behavior diversity is an attribute of most free human societies. Within the time scale of cultural evolution studied here, it is likely that some individuals may become "attached" to their behavior, perhaps as a result of genetic predisposition, or as a result of their beliefs, perhaps simply because they respond too slowly to external stimuli to change. In any of these cases, we may be confronted with the obstinacy that characterizes *PA*s. Interestingly, whenever social diversity is modeled by means of heterogeneous networks of interactions, it can be shown that the most influential individuals are the most connected and the first to adopt cooperative behaviors (Santos & Pacheco, 2006; Santos, et al., 2008), and remain resilient to changes from then on by comparison with the rest of the population. Moreover, the role of the influential person in the overall outcome of evolution is enhanced by their central position, as they efficiently influence a high number of individuals. Hence, the obstinacy of *PA*s may be further amplified as a result of differences in social positions, whenever their location is central in the social network.

In the context of indirect reciprocity and moral systems (Nowak & Sigmund, 1998; Ohtsuki & Iwasa, 2004; Pacheco, Santos, & Chalub, 2006), a behavior somewhat paralleling pathological altruism has been called a phenotypic handicap (Lotem, Fishman, & Stone, 1999), in the sense that it also induces a (more modest) emergence of cooperators. In that case the handicap was associated with defection — defectors would help to stimulate discrimination in the community — whereas in the present model the immutable phenotype is associated with altruism. In both cases, however, one can view this immutability of behavior as maladaptative, which by no means implies that this type of individuals is rare. The result for *PA*s, as shown here, is that the appearance of a single such individual may have a spawning effect in the emergence of cooperation, a feature that would be unavailable until its appearance. The consequences can be devastating for defectors, as we have shown.

A related issue that remains to be investigated is what happens if such an obstinant maladaptation would occur with a defector, instead of a cooperator. The upshot is that the model could then be extended to incorporate pathological defectors ("cheaters," or "psychopaths"), in the population. In a nutshell, pathological cheaters would act to increase the strength of natural selection towards defection. In the simultaneous presence of both pathological altruists and pathological cheaters, the latter would, at most, reduce the fraction of cooperators who optimally coexist with defectors in the population. But these pathological defectors would not be able to counteract the fundamental rift in symmetry introduced by pathological altruists, who would always open a window of viability for cooperation to be maintained in populations.

To summarize, pathological altruists — obstinate cooperators who never change their behavior towards others — introduce profound changes in the evolutionary dynamics of tight communities. In their presence, cheaters no longer push cooperators to extinction. Instead, the population evolves towards *a coexistence* of altruists and cheaters which characterizes its composition most of the time. Ironically, in the currency of cooperation, pathological altruists are very effective in allowing the population to avoid falling into the "tragedy of the commons" doomsday scenario referred to in the beginning. This is done by securing the maintenance of cooperators in the population. In doing so, opportunity is also provided for cheaters to have cooperators to exploit.

References

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390-1396.

Dunbar, R. I. M. (2003). The Social Brain: Mind, Language, and Society in Evolutionary Perspective *Annu. Rev. Anthropol., 32*(1), 163-181.

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

Lotem, A., Fishman, M. A., & Stone, L. (1999). Evolution of cooperation between individuals. *Nature, 400*, 226-227.

Macy, M. W., & Flache, A. (2002). Learning dynamics in social dilemmas. *Proc Natl Acad Sci U S A, 99 Suppl 3*, 7229-7236.

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

McNamara, J. M., Barta, Z., Fromhage, L., & Houston, A. I. (2008). The coevolution of choosiness and cooperation. *Nature, 451*(7175), 189-192.

McNamara, J. M., Barta, Z., & Houston, A. I. (2004). Variation in behaviour promotes cooperation in the Prisoner's Dilemma game. *Nature, 428*(6984), 745-748.

Nowak, M. A. (2006a). *Evolutionary Dynamics*: Belknap/Harvard.

Nowak, M. A. (2006b). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560-1563.

Nowak, M. A., & Sigmund, K. (1998). The evolution of indirect reciprocity by image scoring. *Nature, 393*, 573–577.

Ohtsuki, H., & Iwasa, Y. (2004). How should we define goodness?--reputation dynamics in indirect reciprocity. *J Theor Biol, 231*(1), 107-120.

Ohtsuki, H., Nowak, M. A., & Pacheco, J. M. (2007). Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Physical Review Letters, 98*, 108106

Ohtsuki, H., Pacheco, J. M., & Nowak, M. A. (2007). Evolutionary Graph Theory: Breaking the symmetry between interaction and replacement. *Journal of Theoretical Biology, 246*(4), 681-694.

Pacheco, J. M., Santos, F. C., & Chalub, F. A. (2006). Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comput Biol, 2*(12), e178.

Santos, F. C., & Pacheco, J. M. (2006). A new route to the evolution of cooperation. *J Evol Biol, 19*(3), 726-733.

Santos, F. C., Santos, M. D., & Pacheco, J. M. (2008). Social diversity promotes the emergence of cooperation in public goods games. *Nature, 454*(7201), 213-216.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*: Cambridge University Press.

Szabó, G., & Tőke, C. (1998). Evolutionary prisoner's dilemma game on a square lattice. *Physical Review E, 58*(1), 69-73.

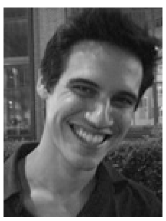
Taylor, P. D., Day, T., & Wild, G. (2007). Evolution of cooperation in a finite homogeneous graph. *Nature, 447*(7143), 469-472.

Traulsen, A., Nowak, M. A., & Pacheco, J. M. (2006). Stochastic Dynamics of Invasion and Fixation. *Phys. Rev. E, 74* 011090.

Van Segbroeck, S., Santos, F. C., Lenaerts, T., & Pacheco, J. M. (2009). Reacting differently to adverse ties promotes cooperation in social networks. *Physical Review Letters, 102*, 058105.

About the authors

Jorge M. Pacheco (Oporto, 1958) is a Full Professor at the Department of Mathematics of the University of Minho (Portugal). He studied Theoretical Physics at the University of Coimbra (Portugal) and obtained his Ph.D. at the Niels Bohr Institute (Denmark, 1990). He is active in a variety of research topics, ranging from quantum many-body physics to the mathematical description of complex processes such as somatic evolution of cancer, the evolution of cooperation, the evolution of culture and complex network science.
(email:pacheco@cii.fc.ul.pt
webpage: http://sites.google.com/site/jorgempacheco/).



Francisco C. Santos (Lisbon, 1981) studied Physics at the University of Lisbon (Portugal) and earned his Ph.D. in Computer Science at the Free University of Brussels (Belgium, 2007). After 3 years as associate researcher in Brussels, he is currently an associate researcher of the New University of Lisbon (Portugal). His interests span several aspects of complex adaptive systems, from the structure of social networks and the dynamics of human cooperation, to the evolution of social norms and reputation-based systems.
(email : xsantos@gmail.com;
webpage : http://iridia.ulb.ac.be/~fsantos/).