Adaptive parsimony as an evolutionary solution to the equilibrium selection problem

Jean-Baptiste André[1]

[1]Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, PSL Research University, CNRS, Paris France

June 28, 2023

Abstract

Many games, especially repeated games, have multiple Nash equilibria. This multiplicity limits the predictive power of game theory for understanding animal behavior, and plays an important role in evolutionary anthropology as a seemingly irrefutable argument that cultural group selection is necessary for human cooperation. In this article, I propose a solution to this problem inspired by the notion of convergence stability from adaptive dynamics. The multiplicity of equilibria is due to the possibility of strategies that are arbitrary in the sense that they are individually adaptive only because others use them. While arbitrariness is a possibility in standard game theory, since it is always possible to design patterns of behavior as complex as one wishes, it cannot be gradually shaped by biological evolution. Evolution can stay in an arbitrary equilibrium if it starts there, but it cannot converge to an arbitrary equilibrium if it starts from a different initial state. I propose an equilibrium refinement, the concept of evolutionarily parsimonious equilibrium, that captures this convergence constraint by relying on the premise that a complex strategy cannot evolve as a by-product of selection for a simple strategy. Using examples, I show that this refinement supports the selection of biologically reasonable equilibria in many of the most important games in the literature. In particular, it eliminates the vast majority of equilibria in repeated games and provides a better understanding of the conditions necessary for the evolution of reciprocal cooperation.

1 Introduction

Many of the most important games for understanding social behavior have multiple Nash equilibria. This is the case, for example, of the ultimatum game (Binmore and Samuelson 1994), of most signaling games (Riley 2001; Sobel 2012), of bargaining games (Binmore

1985; Nash 1950; Rubinstein 1982; Shaked 1986), and, most famously, of repeated games, for which this multiplicity is called the folk theorem (Aumann and Shapley 1994; Boyd 2006). To illustrate, in the pairwise repeated prisoner's dilemma, any quantitative level of cooperation can be supported by a Nash equilibrium, from no cooperation to maximum cooperation, through all possible intermediates. This indeterminacy limits the predictive power of game theory for understanding animal behavior.

To work around this limitation, game theoreticians, as well as some evolutionnary biologists, have developed refinements of the Nash equilibrium concept. They have sought ways to reduce the multiplicity of equilibria by arguing that some equilibria are more "reasonable" from a rational point of view than others, or are more likely to be the result of evolution.

In standard, non-evolutionary game theory, the most prominent of these refinements is the concept of subgame perfection (Selten 1965, 1973, 1975). It captures the idea that some Nash equilibria are implausible because they involve non-credible threats. For example, in the ultimatum game, any strategy profile that consists of the proposer offering $x \in]0, 1]$ and the receiver refusing any offer strictly less than x is a Nash equilibrium, but not a subgame perfect equilibrium, because it involves the threat that if the proposer deviates from the equilibrium, then the receiver would behave irrationally by refusing a strictly positive offer. The only subgame perfect equilibrium in the ultimatum game is for the proposer to offer exactly $x = 0$, which is indeed the most reasonable outcome, consistent with the strategic asymmetry of the game.

However, for evolutionists interested in understanding how evolution has shaped the social behavior of humans and other animals, the application of the concept of subgame perfection is not straightforward. Evolution by natural selection can lead to subgame perfect equilibria under certain assumptions (e.g., assuming errors in the course of the game; Selten 1983, 1988), but it need not always lead to such equilibria (Gale et al. 1995). Moreover, the concept of subgame perfection does not reduce the diversity of equilibria in important cases. It does not reduce the diversity of equilibria in games without subgames. And it does not reduce the diversity of equilibria in repeated games, which is one of the most important challenges to understanding cooperation in the human species. In a game repeated long enough, all possible levels of cooperation can be supported by a subgame perfect equilibrium (Aumann and Shapley 1994). The diversity of subgame perfect equilibria is thus as great in practice as the diversity of Nash equilibria.

Based on evolutionary considerations, two other mechanisms have been proposed to refine the predictions of game theory beyond the Nash equilibrium concept.

The first of these mechanisms, proposed by some evolutionary anthropologists and economists, is based on cultural diversity and group selection. The idea is that different human groups will culturally reach different Nash equilibria, and that competition between groups will secondarily favor those groups that have reached the more socially efficient equilibrium, leading to the spread of "cultural norms" of cooperation. This mechanism has been very influential among scholars interested in evolutionary approaches to human behavior. It has been proposed and somewhat formalized by a school of biological anthropologists who have made it the core of the latest version of their theory called “cultural group selection” (Boyd and Richerson 2009). And it has also been proposed, albeit verbally, in the work of game theorists and philosophers, such as in Binmore’s theory of natural justice (Binmore 2009) or Gintis’s theory of property (Gintis 2007).

The second of these mechanisms uses a famous refinement of the Nash equilibrium concept from biology: the concept of an evolutionarily stable strategy (hereafter ESS). An ESS is defined by two conditions. The first corresponds to the Nash equilibrium condition. The second, called Maynard-Smith’s condition, refines the Nash concept by taking into account the payoffs of rare mutants that interact with each other. Somewhat paradoxically, in the vast majority of games studied by biologists, this second condition plays no role at all, because these games have *strict* Nash equilibria, in which case the second condition does not matter. The main exception is dynamic games. These games generally do not have strict Nash equilibria, so the second ESS condition plays a role and eliminates equilibria that are not socially efficient. For example, in the repeated prisoner’s dilemma, all levels of cooperation can be supported by a Nash equilibrium, but only the maximum possible level of cooperation is ESS (André and Day 2007; Binmore and Samuelson 1992; Fudenberg and Maskin 1990; Nowak, Sasaki, et al. 2004). A variant of this mechanism, which tends to favor less cooperative equilibria, is to consider the effects of interaction between two different mutants, a neutral mutant on the one hand, and another mutant that can use the first as a stepping stone to destabilize the resident (Boyd and Lorberbaum 1987; García and Van Veelen 2016; Van Veelen 2012; Van Veelen and Garcıa 2019).

Both mechanisms–group selection and mutant-mutant interaction–are in principle valid from an evolutionary point of view, but the conditions necessary for them to be effective are very restrictive. The first requires the existence of structured groups that have the opportunity to establish a diversity of social norms and then compete with each other. The second relies on selection due to the presence of mutants, which is a weak selective force–because it is second order in the effect of mutations and proportional to the frequency of mutants–that can play a role only if all other forces are very weak.

In this article, I argue that neither mechanism is necessary because there is another, more robust and parsimonious constraint that selects equilibria in biology: the constraint of evolutionary convergence.

Contrary to what the success of the ESS concept might suggest, the most important conceptual innovation in evolutionary biology with respect to the Nash equilibrium is not the second ESS condition, but the concept of stability by convergence developed for the study of quantitative traits (Eshel 1983; Eshel et al. 1997; Geritz et al. 1998). Stability by convergence captures the fact that evolution is a gradual process. It occurs through a succession of fixations of adaptive mutations of relatively small effect, each of which must be advantageous at the time of its appearance. Some strategies are Nash equilibria and ESSes, but will not evolve by natural selection because they cannot be achieved by such a gradual process. They are then said to be *evolutionarily* stable but not *convergence* stable.

In the specific case of quantitative traits, the idea that adaptive mutations have small effects and that adaptation is gradual has a very clear mathematical translation. It is much less straightforward to understand how this notion translates in the case of games, such as repeated games, whose strategy sets are typically discrete rather than quantitative. As a result, the notion of stability by convergence has never been used or formalized in dynamic games.

Yet the notion that adaptive mutations are of weak effect, and the ensuing notion that the adaptive process is gradual, is not a mere technicality specific to the evolution of quantitative traits. On the contrary, it captures a fundamental principle of biological evolution. The very problem with complex adaptations is that they are organized and thus *improbable* states (Dawkins 1996). As a result, the probability that they could arise by chance, from one or a few macro-mutations, is by definition extremely small–Fisher (1930)'s geometric model of adaptation can be seen as a formalization of this idea. Natural selection is a major scientific breakthrough because it solves the apparent puzzle of adaptation by breaking down an unlikely random occurrence of complex order into an accumulation of a series of small effect, and therefore plausible, adaptive mutations that progressively increase the amount of functionality and order in biological systems. The convergence condition is therefore a fundamental and general constraint on biological evolution.

This constraint dramatically reduces the range of strategies that can be produced by evolution. The multiplicity of equilibria arises from the fact that many games have a circular dimension (Boyd 2006). What is adaptive for one individual depends on what others do, and vice versa, opening up the possibility of an infinite number of equilibria. If other individuals cooperate only on odd days, then it is adaptive to do the same. If they cooperate only on rainy days, then it is pointless to try to cooperate on sunny days, and so on. These strategies imply conditional abilities–e.g., cooperating only on odd days–that have no adaptive value in themselves, but are valuable only if other individuals have the same abilities. Such abilities may well be part of a *equilibrium*, but the circular nature of their adaptive advantage makes it impossible for evolution to *converge* to them through the gradual accumulation of weak-effect mutations.

Thus, not all equilibria can be reached by biological evolution. Some equilibria involve circular capabilities that have no value per se, but only value if others have them, while others do not involve such capabilities. Only the latter can be attained through the gradual process of evolutionary adaptation.

In this article, I propose an equilibrium concept, the concept of evolutionarily parsimonious equilibrium, that allows to capture this convergence constraint by relying on the premise that a complex strategy cannot evolve as a by-product of selection for a simple strategy. Using examples, I show that this concept supports the selection of biologically reasonable equilibria in many of the most important games in the literature. In particular, it eliminates the vast majority of equilibria in repeated games and provides a better understanding of the conditions necessary for the evolution of reciprocal cooperation.

2 Evolutionarily parsimonious equilibrium

First, a restricted formulation of the concept is presented to help understand its logic. The general formulation follows.

Let G be an extensive-form game, defined by a finite set of n players, a game tree representing the sequence of player decisions, a set of final nodes with payoff values for each player, and a partition of the non-terminal nodes into $n+1$ subsets, one for each player and one for Nature's moves. Each node of the game tree is called a "history". We say that a history belongs to a player if he is the one making a decision at that history. A player's strategy is then defined by a partition of his histories and a mapping from the elements of this partition to the set of possible behavioral policies.

In the restricted version of the analysis, I consider a symmetric game in which, for exogenous reasons beyond the players' control, each player has only 2 different information states. In other words, of all their histories, players can distinguish at most two types, and no more.

The goal of the analysis is to determine the conditions necessary for evolution by natural selection to drive the transition from a strategy in which players do not distinguish between the two information states and play the same policy in each history of the game, to a strategy with an additional level of contingency in which players play two different policies in each of the two information states.

Consider an ancestral strategy S that does not distinguish between the two states. In pure strategy, this means that players unconditionally play exactly the same action in all their histories. In mixed strategy, this means that they unconditionally draw their actions from the same probability distribution or from infinitesimally close probability distributions in all histories.

Consider another strategy, S', that distinguishes between the two information states and plays policies b_1 and b_2 in each state where at least one (or both) differs from b.

Then let two other strategies S_1 and S_2 be defined as follows: like S they do not discriminate between the two states, but unlike S they unconditionally play the policies b_1 and b_2 in both states.

We call $P(X,Y)$ the payoff of a player with strategy X interacting with a player with strategy Y.

Proposition: The evolutionary transition from strategy S to strategy S' is possible if the following three conditions are satisfied:

$$P(S', S) > P(S, S) \tag{1}$$

$$P(S', S) > P(S_1, S) \tag{2}$$

$$P(S', S) > P(S_2, S) \tag{3}$$

In this case, we say that S' is *parsimoniously evolvable* from S.

Condition 1 requires that S' is strictly better than S. A neutral strategy with a superfluous conditional ability is not parsimoniously evolvable from an unconditional resident.This captures the idea that strategies cannot become more complex, i.e., acquire novel behavioral dispositions, without selection pressures to favor it. The other two conditions are more subtle. They require that neither of the two simpler strategies that would play the same policy in both states is able to do better than S' against the resident. In other words, S' is parsimoniously evolvable from S, provided that its extra complexity really gives it an advantage over strategies without that complexity.

2.1 General formulation

Consider the same game G as defined above, but this time not necessarily symmetric, and with no limit on the number of information states.

Define a symmetric positive function $d(X, Y)$ that measures the mutational distance between any two strategies X and Y. This mutational distance is intended to measure a transition probability (see below). The greater the mutational distance between two strategies, the less likely it is that a random biochemical event will cause a lineage to switch from one to the other.

Consider a resident strategy S and any alternative strategy S'. We say that S' is parsimoniously evolvable from S in one step if the following condition is satisfied:

$$\forall S'' \text{ with } d(S'', S) < d(S', S),\ P(S', S) > P(S'', S) \tag{4}$$

i.e. S' is strictly better against S than any strategy whose distance to S is smaller. In particular, this implies that S' is strictly better against S than S itself.

This can then be generalized to multiple steps. The evolutionary transition from an ancestral strategy S to a derived strategy S' is possible if there exists an ordered list of strategies of finite length, where S is the first one in the list, S' is the last one in the list, and each element of the ordered list is parsimoniously evolvable in one step from the previous element. In this case, we say that S' is parsimoniously evolvable from S.

Finally, the concept of an evolutionarily parsimonious equilibrium is defined from an ancestral strategy S_0. We say that a strategy S^* is an *evolutionarily parsimonious equilibrium* from S_0 if it is a Nash equilibrium and it is parsimoniously evolvable from S_0. Note that if S_0 itself is a Nash equilibrium, then the only evolutionarily parsimonious equilibrium from

S_0 is S_0 itself.

The concept of evolutionarily parsimonious equilibrium is independent of the exact choice of the mutational distance function $d(\cdot)$. In principle, any distance measure could be considered. However, the biological intuition that this approach aims to capture is that conditional abilities are adaptations that can only arise if they are favored as such. Thus, it is essential that the distance measure between two strategies take into account their conditional capacities (as was done in the restricted approach above). In the examples shown in the paper, I measure mutational distances using finite automata. The distance between two strategies is computed in 3 steps (see Appendix A for details): (1) I build the finite automata needed to implement each strategy. (2) I search for the best possible alignment between the two automata. (3) I comprehensively count the number of differences between the two automata once they are aligned.

2.2 Evolutionary interpretation

The concept of an evolutionarily parsimonious equilibrium attempts to capture in a simple way an underlying biological hypothesis: adaptive mutations are all the rarer the stronger their effect. This hypothesis, hereafter referred to as the relative rarity of adaptive mutations hypothesis, is standard in evolutionary biology and was captured geometrically by Fisher (1930).

An adaptive mutation that transforms a strategy S into a strategy S' at a mutational distance k away (i.e., with $d(S', S) = k$) is rarer and thus occurs later in expectation than all mutations that transform S into S'' at a smaller mutational distance (i.e., with $d(S'', S) < k$). Thus, if one of these closer mutants can invade a population of S, it will do so before the more distant mutant has a chance to rise. As a result, the one-step evolutionary transition from S to S' cannot occur.

The consequence of the relative rarity of adaptive mutations hypothesis is that a complex behavior cannot evolve as a by-product of selection in favor of a simpler behavior. For example, a conditional behavior cannot evolve as a by-product of selection for a constitutive behavior. It can only evolve if it provides an advantage over the constitutive behavior.

Note that according to the relative rarity of adaptive mutations hypothesis, one should require the following stricter conditions for S' to be evolvable from S in one step:

$$P(S', S) > P(S, S) \tag{5}$$

$$\forall S'' \text{ with } d(S'', S) < d(S', S),\ P(S'', S) \leq P(S, S) \tag{6}$$

These stricter conditions require (1) that S' is strictly better against the resident, and (2) that there is no mutant strategy S'' that is both more similar to the resident and also strictly better against the resident. In other words, under these stricter conditions, even if S' is better against the resident than all closer strategies, S' cannot be considered evolvable in one step if some closer strategies are able to invade the resident before it.

It is not entirely clear which of these two conditions is the better choice (condition 4 or the stricter condition above). The strict condition is a more natural consequence of the relative rarity of adaptive mutations hypothesis, but it is also stricter. To show that the present approach significantly reduces the multiplicity of equilibria, the less restrictive condition is preferable because it is more conservative. Another way of saying this is that condition 4 is a *necessary* condition, not a sufficient condition, for a strategy to be able to evolve by natural selection. Note that in practice, in all the examples we will consider in this paper, both conditions have exactly the same consequences.

3 Application to examples

3.1 Divide-the-dollar game

The divide-the-dollar game is a two-player sharing game. Each player must announce in advance the fraction of a resource he is claiming, without knowing what the other player is claiming. Each player gets what he asked for if the sum of the two requests does not exceed the total amount of the resource, otherwise both players get nothing. Here I consider the simple case where the two players are initially symmetric, and I consider only the set of pure strategies.

I assume that the game starts with a move by Nature (i.e., an event that cannot be controlled by the players) that creates an asymmetry between the two players, e.g., one player arrives from the right, the other from the left, one player arrives before the other, and so on.

If the symmetry breaking is physically undetectable to both players, then they cannot condition their behavior on it. In this case, in pure strategy, both players must make exactly the same demand in equilibrium, and the only possible Nash equilibrium of the game is for the players to demand exactly half of the resource, i.e., $x^* = 0.5$. Any demand other than $x = 0.5$ is not in equilibrium, as players would have an incentive to deviate (things are a bit more complex when considering mixed strategies, see Skyrms 1996).

On the other hand, if the players have the sensory capacity to detect the symmetry breaking, they can condition their demands on it, dramatically expanding the range of possible equilibria. Suppose one individual is in role A and the other in role B after Nature's move, then all pairs of demands (x^*_A, x^*_B) with $x^*_A + x^*_B = 1$ are Nash equilibria in pure strategy, since neither player gains anything by deviating. Moreover, all these equilibria are strict Nash (deviations have strictly positive costs) and therefore evolutionarily stable.

Thus, the Nash equilibrium is unique and symmetric only in the very specific situation where the players are in a completely identical state with respect to their perceptual abilities. In practice, since it's very unlikely that two individuals will be in completely symmetric perceptual states, there will always be an infinite number of Nash equilibria.

Biologically, however, it seems unrealistic that complete *sensory* symmetry is necessary for evolution to lead to a symmetric outcome. It would seem more natural for a notion of *ecological* symmetry to play a role. The concept of evolutionarily parsimonious equilibrium captures this intuition.

Consider an ancestral strategy S that does not distinguish between the two states and requests the same $q \in [0,1]$ regardless of which state, A or B, it is in. Consider also a conditional strategy S' that requests q_A when in state A and $q_B \neq q_A$ when in state B. And finally, consider two unconditional strategies S_A and S_B that request q_A and q_B, respectively, regardless of their state, i.e., regardless of the situation A or B in which they find themselves. We will show that S' can never be parsimoniously evolvable from S in one step (the restricted formulation of evolutionary parsimony can be used here because individuals can be in only two information states).

The first condition for S' to be parsimoniously evolvable from S in one step is that S' is strictly better than S against S itself (condition 1). This implies that S', at least in one of the two states $i \in \{A, B\}$, makes a request $q_i \in R$, where $R =]q, 1-q]$ if $q < 0.5$ and $R =]0, 1-q]$ if $q > 0.5$. Since S' by definition makes a different request in the other state $j \neq i$, one of its two requests is necessarily strictly better than the other. If $q_j \notin R$, then q_i is strictly better than q_j, and if $q_j \in R$, then the larger of the two requests is strictly better than the other. So there will always be one of the two unconditional strategies, S_A or S_B, that is strictly better than S' against the resident, i.e. $\forall (q_A, q_B)$, $P(S_A, S) > P(S', S)$ or $P(S_B, S) > P(S', S)$. At least one of the two conditions 2 or 3 is not true. Therefore, S' is not parsimoniously evolvable from S in one step.

Extending the reasoning to multiple evolutionary steps, a strategy that discriminates between the two states can never be parsimoniously evolvable from an ancestral strategy that does not. That is, from any ancestral unconditional strategy, the only evolutionarily parsimonious equilibrium is the symmetric strategy that demands $q^* = 0.5$ in both states.

All asymmetric equilibria are excluded by the parsimony condition, because they require that individuals have the ability to condition their actions on their state, and the only selection pressure that can lead to the evolution of such an ability comes from the pre-existence of the same ability in other players. Such circularity allows asymmetric strategies to be evolutionarily stable, but it does not allow evolution to converge on them.

3.2 Hawk-Dove game

This second example will be contrasted with the first to understand the conditions under which evolution can produce a symmetry breaking.

Let us also consider two players, previously distinguished by a move by Nature, playing a Hawk-Dove game with the payoff given in Table 1. Here we consider the full set of mixed strategies. That is, the strategy played by each individual is characterized by a pair (p_A, p_B), where each $p_i \in [0,1]$ is the probability of playing *Hawk* when in state i.

Consider an ancestral strategy (p, p) that plays *Hawk* with the same probability $p \in [0,1]$

	Hawk	Dove
Hawk	P	T
Dove	S	R

Table 1: Payoff matrix of the Hawk-Dove game. Parameters must respect $T > R > S > P$

regardless of its state. By the same reasoning as in the divide-the-dollar game above, since this strategy behaves identically regardless of its state, there can be no selection pressure to condition one's action on one's state when playing against it. The fixation of parsimonious adaptive mutations from this ancestral strategy can therefore only lead to a strategy that is itself unconditional. Thus, evolution converges to a strategy profile where both players play *Hawk* with an intermediate probability given by $p_A = p_B = p^* = (T - R)/(T + S - P - R)$.

The difference with the divide-the-dollar game is what can happen once this situation is reached. Consider the effect of introducing an uncontrollable infinitesimal disparity in the respective probabilities of playing *Hawk* in states A and B. More specifically, suppose that for some reason all players in state A tend to play *Hawk* a little more often than players in state B, i.e., we have $p_A = p^* + \delta$ and $p_B = p^* - \delta$, where $\delta \ll 1$ and $\delta > 0$.

Now consider a mutant strategy that is able to condition its behavior on its state, so that it plays *Hawk* with probability $p_A > p^*$ when in state A, and $p_B < p^*$ when in state B. We want to find out if this mutant is parsimoniously evolvable from the resident (p^*, p^*) in one step.

To the first order in δ, the net advantage of the mutant over the resident is $\frac{1}{2}(T + S - P - R)(p_A - p_B)\delta$, which is strictly positive if $p_A > p_B$. Thus the condition 1 is satisfied.

Since the strategy (p^*, p^*) is defined by the fact that playing *Hawk* or *Dove* against it is strictly neutral, the two unconditional mutants playing (p_A, p_A) and (p_B, p_B) are themselves neutral against (p^*, p^*). Therefore, the conditional strategy (p_A, p_B) has a net advantage $\frac{1}{2}(T + S - P - R)(p_A - p_B)\delta > 0$ over them. So both conditions 2 and 3 are fulfilled.

In sum, the conditional strategy (p_A, p_B) is parsimoniously evolvable from the ancestral strategy (p^*, p^*) as long as there is some infinitesimal difference in the probability that the ancestral strategy plays *Hawk* on each state. Once symmetry breaking is achieved, the resulting evolutionarily parsimonious equilibrium is a strategy in which players fully condition their behavior on their state. That is, two strategy profiles are evolutionarily parsimonious equilibria: $(1, 0)$ and $(0, 1)$, corresponding to the so-called Bourgeois strategy from evolutionary biology (Maynard Smith and Parker 1976) and to the more general notion of correlated equilibrium in game theory (Aumann 1974; Aumann 1998).

The emergence of an arbitrary asymmetry from a symmetric ancestral strategy is possible here because even an infinitesimally small initial asymmetry is sufficient to select for a strict, non-infinitesimal, divergence between the two states. In other words, selection will amplify even the smallest difference between players, allowing evolutionary convergence toward a state-dependent strategy.

In the divide-the-dollar game considered earlier, on the other hand, an initial infinitesimal difference between players A and B could only select for the same infinitesimal difference, not for an increase in that difference. Thus, there could be no evolutionary path from initially state-independent to state-dependent strategies.

Intuitively, role specialization is a "reasonable" evolutionary outcome in a game where coordination is useful because it allows players to solve an actual problem they face, as in the hawk-dove game (or other coordination games). In contrast, role specialization is only an unnecessary complication in games where coordination has no functional use, such as when sharing a divisible resource, since it would only add unnecessary complexity to the players' decision mechanism. The concept of adaptive parsimony accounts for this intuitive difference between the two types of situations.

3.3 Ultimatum game

Consider two asymmetric players, one playing the role of proposer and the other playing the role of responder, who need to share a resource of size 1. To do this, the proposer first offers to give a fraction p of the resource to the responder, who then decides whether to accept or reject the offer. If the offer is accepted, the proposer and the responder receive $1-p$ and p respectively, otherwise both players receive 0. The proposer's strategy is defined by an offer $p \in [0,1]$, while the responder's strategy is a mapping from the set $[0,1]$ to the set $\{\text{accept}, \text{reject}\}$.

All pairs of strategies where the proposer offers $p^* \in [0,1]$ and the responder accepts p^* and rejects any $p < p^*$ are Nash equilibrium profiles, since no deviation is advantageous. Among these equilibria, however, the equilibrium profile in which the proposer offers $p^* = 0$ and the responder accepts any offer, no matter how small, is more "reasonable" than the others because it is consistent with the strategic asymmetry of this game.

This intuition that one particular equilibrium is more reasonable than others is well captured by the concept of subgame perfection. Any equilibrium in which the responder rejects some strictly positive offers implies that he behaves suboptimally in all subgames in which these offers are made, and is therefore not subgame perfect.

The concept of adaptive parsimony captures the same intuition in a more evolutionarily relevant way. The responder's acceptance of some offers and rejection of others involves a conditional decision mechanism. He must be able to evaluate the offers made to him and make a decision based on their value. In other words, any strategy that consists of the responder accepting some offers and rejecting others contains more conditionality than a strategy that consists of accepting or rejecting all offers. The question, then, is whether a responder's strategy with conditional acceptance is parsimoniously evolvable from an ancestral strategy without such conditionality.

First, consider an ancestral strategy profile where the proposer offers $p_0 \in [0, 1]$ and the responder accepts all offers unconditionally. Then consider a mutant responder that accepts offers only if they are strictly above a threshold $x > 0$. The mutational distance between this mutant and the resident is strictly positive, while the mutant is at best neutral (if $x < p_0$) and at worst counter-selected (if $x \geq p_0$). Therefore, condition 4 cannot be satisfied.

Now consider an ancestral strategy profile S where the proposer offers $p_0 > 0$ and the responder unconditionally rejects all offers. Then consider a conditional responder strategy S' that consists of accepting offers only if they are strictly above a threshold $x \in]0, 1[$. And consider a third responder strategy S'', which consists of accepting all offers unconditionally.

If $x < p_0$, then we have $P(S', S) > P(S, S)$, i.e. the conditional acceptor mutant S' is strictly favored by selection in a resident population of S, because it accepts the offer p_0 instead of rejecting it. However, the unconditional mutant S'' is just as good as S', i.e. $P(S'', S) = P(S', S)$. In terms of mutational distance, we have $d(S', S) > d(S'', S)$, because S' has both one additional state and two different transitions (see Appendix A and Figure 1), while S'' only plays a different action in its single state. Therefore, condition 4 cannot be satisfied.

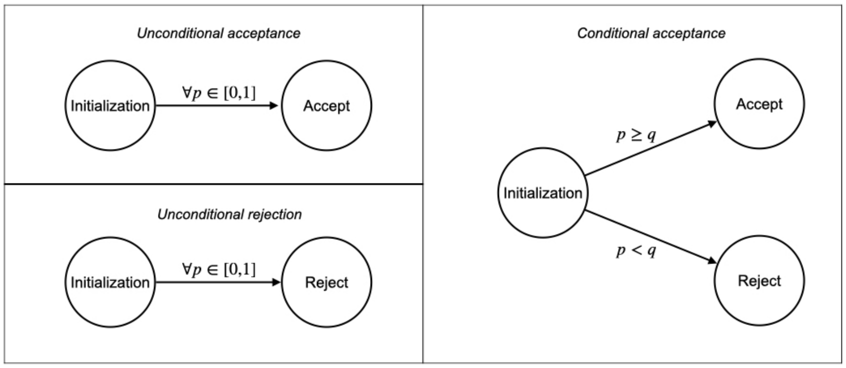

Figure 1: **Strategies in the ultimatum game represented as finite automata**. The distance between unconditional acceptance and unconditional rejection is $\Delta_B = 1$. The distance between either unconditional acceptance or unconditional rejection and conditional acceptance is $\Delta_E + \Delta_T = 1 + 2$.

From any unconditional ancestral strategy profile where the proposer offers $p_0 > 0$, the only evolutionarily parsimonious equilibrium is where the proposer offers the minimum possible offer $p_0 \approx 0$ and the responder accepts any offer.

One way to see this is to recognize that in the absence of reciprocity or genetic relatedness, there is never any selection pressure on individuals to be able to *reject* useful resources available in their environment. The only selection pressure on individuals is to *exploit* as many resources as possible, whether or not they originally came from another individual who gave them away in the first place.

3.4 Repeated prisoner's dilemma

Consider a discrete repeated prisoner's dilemma with two symmetric players who have the choice at each stage of the game between cooperating (C) and defecting (D), and consider the family of strategies called Grim, which consists in responding to all deviations, including deviations by the player herself, with an infinite series of defections.

Let us start with the simplest strategy of this family, here called Grim_1, which consists of the following behavioral rule: Play C at every stage of the game, if you and your partner have always played according to Grim_1 in the past, otherwise play D (see Figure 2 for a representation of this strategy and other strategies in the repeated prisoner's dilemma as finite automata). If the game has a sufficiently high continuation probability, then Grim_1 is a Nash equilibrium and a subgame perfect equilibrium. Deviating from Grim_1 by playing D when no one has ever deviated leads to the end of cooperation, which is costly. Conversely, deviating from Grim_1 by playing C when you should have played D has an immediate cost and no benefit, since your partner will play Grim_1 and thus defect in all future stage games anyway.

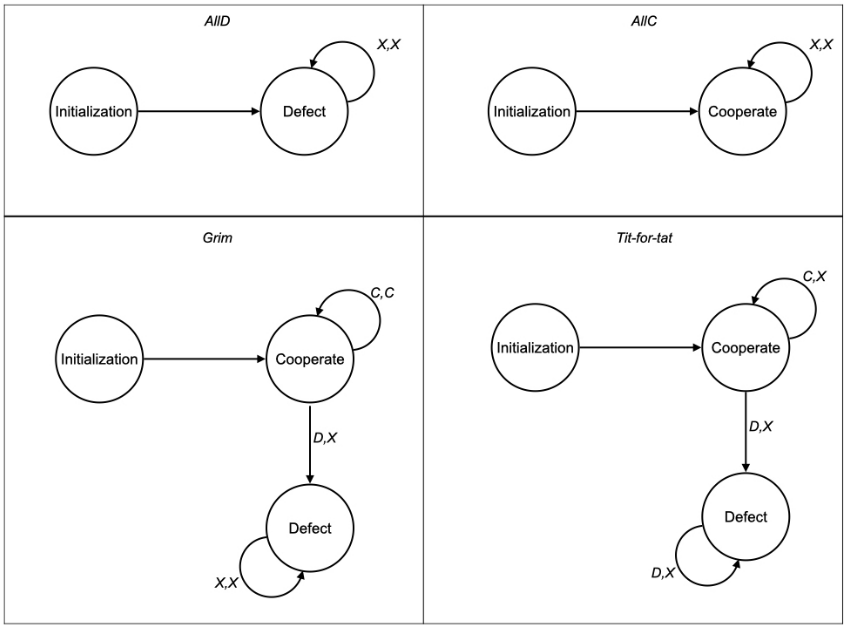

Figure 2: **Strategies in the prisoner's dilemma represented as finite automata**. Mutational distances are as follows: $d(\text{AllC}, \text{AllD}) = 1$, $d(\text{AllC}, \text{TFT}) = d(\text{AllD}, \text{TFT}) = 4$, $d(\text{AllC}, Grim) = 4$.

For the same reason, many other strategies in the Grim family are also Nash and subgame perfect equilibria. For example, consider $\text{Grim}_{\text{even}}$, which consists of unconditionally playing D every odd round and C every even round as long as no one has ever deviated from $\text{Grim}_{\text{even}}$. That is, $\text{Grim}_{\text{even}}$ is a Grim_1 that discounts every odd-stage game. It is also subgame perfect, since no deviation can be beneficial before a $\text{Grim}_{\text{even}}$. By the same logic, any Grim_x that cooperates a fraction $x \in [0, 1]$ of the stage games is also subgame perfect.

In addition, if a prior move by Nature breaks the symmetry between the two players, then any $\text{Grim}_{x,y}$ in which player A cooperates a fraction x of the time and player B cooperates a fraction y of the time is a subgame perfect equilibrium as long as the payoff to both players is greater than what they would gain by not cooperating at all.

An intuitive way to understand this multiplicity of equilibria is to see that there is an element of coordination in reciprocal cooperation (Boyd 2006). To be in equilibrium is to cooperate in the way "required" by one's partner. If the partner requires cooperation only in certain rounds, then cooperation is adaptive in those rounds and not in others.

This diversity, known as the folk theorem, is present in all repeated games. Provided a game is repeated a sufficiently large number of times, all payoff profiles can be supported by a subgame perfect equilibrium if they are individually rational, i.e., if they compensate all players for their outside options.

From a biological point of view, this multiplicity seems like a strange complication. Why should individuals refuse to cooperate in some rounds? Why should natural selection favor strategies of partial cooperation if they only add unnecessary complexity? Again, the concept of evolutionary parsimony captures this biological intuition.

Partial cooperation strategies condition their action on a physically perceptible property of stage games (e.g., round parity), even though this property does not predict the payoff profile of the stage game. The *only* selection pressure that can make such an arbitrary conditionality adaptive comes from the existence of the same conditionality in others. For this reason, a strategy with an arbitrary conditionality cannot be an evolutionarily parsimonious equilibrium from an ancestral strategy without the conditionality.

The problem, however, is that the repeated prisoner's dilemma is not the right model to formalize the consequences of this equilibrium refinement. Starting from an ancestral strategy of pure defection, the only evolutionarily parsimonious equilibrium is actually pure defection itself. Even the simple cooperation strategy Grim_1 is not an evolutionarily parsimonious equilibrium from an unconditional ancestor. Thus, the concept of parsimony eliminates much more than strategies of partial cooperation. It eliminates any possibility of reciprocal cooperation. In fact, this negative result is a mere consequence of the fact that pure defection is a Nash equilibrium in the repeated prisoner's dilemma. By definition, the only evolutionarily parsimonious equilibrium from a Nash equilibrium is this Nash equilibrium itself.

The fact that pure defection is a Nash equilibrium in the repeated prisoner's dilemma is generally not considered a problem for the evolution of reciprocity in this game. Many models in evolutionary game theory suggest that stochasticity may be sufficient to escape this equilibrium and reach more cooperative ones (André and Day 2007; Binmore and Samuelson 1992; Fudenberg and Maskin 1990; Nowak, Sasaki, et al. 2004). The concept of parsimony, on the other hand, attempts to capture the idea that complex biological functions cannot emerge without selective pressures that gradually shape them. In the particular case of the repeated prisoner's dilemma, it implies that, contrary to what these models suggest, the stochastic transition from pure defection to conditional cooperation is actually not possible. It would imply that the ability to cooperate conditionally should emerge in a fully functional form by mutation, as a hopeful monster. I have already formalized elsewhere why this is unlikely (André 2014) and tested it in a simulated robotics model (André and Nolfi 2016). The notion of parsimony formalizes the consequences of this constraint in a more general framework.

By effectively ruling out the possibility of reciprocity in the repeated prisoner's dilemma, the concept of evolutionarily parsimonious equilibrium provides formal support for a long-standing criticism of this game as a model of reciprocal cooperation. Empirically-minded authors have long pointed out that while it is easy to understand the existence of selection pressures on the ability to choose the best partners in partner choice games, the existence of selection pressures that might have shaped the ability to cooperate reciprocally in the repeated prisoner's dilemma is much less obvious (Bergmuller et al. 2007; Bshary et al. 2016; Connor 1995; Noë and Hammerstein 1994; Roberts et al. 2021). The concept of parsimony captures their informal argument. The repeated prisoner's dilemma is not the right game for understanding the evolution of reciprocal cooperation.

3.5 Investment game: from variable by-products to reciprocity

Reciprocal cooperation can be supported by an evolutionarily parsimonious equilibrium, provided that we consider a game in which individuals are directly under selection to cooperate conditionally, independent of the prior existence of such conditionality in others. We can illustrate this in a simple game that I will briefly describe here and analyze in more detail in Appendix B (see also Geoffroy and J. B. André 2021).

Consider a repeated game between two individuals, called the investor and the responder. The game is preceded by a move by Nature that determines the type of stage game that the individuals will then play repeatedly. Three types of stage games are possible: (1) The stage game can be a by-product cooperation game, in which both the responder and the investor automatically benefit from the investment. (2) The stage game can be a trust game, in which the investment has a net cost to the investor, but the responder can reward him afterwards and still make a net profit. (3) The stage game can be an inefficient cooperation game, in which the investment has a net cost to the investor and the responder cannot reward him in a mutually beneficial way.

Regardless of the game type chosen by Nature, the investor has two possible actions in each round: *Invest* or *Decline*. Then, if the investor decides to invest in a given round, the responder has two possible actions: *Reward* or *Defect*. If the investor chooses to decline, then the responder has no decision to make.

I assume that only the responder knows the type of game being played, which captures the essence of what makes reputation relevant. The responder is informed about his own payoffs, and the investor can only discover them indirectly by observing the responder's behavior.

Consider a strategy profile (I^*, R^*), where (i) the investor tests each game for one round and continues to invest in subsequent rounds only if there is a net gain, and (ii) the responder rewards the investor only if the game is a trust game, i.e., only if rewarding is useful to ensure the investor's continued participation (see Figure 3 for a representation of this strategy as a finite automaton).

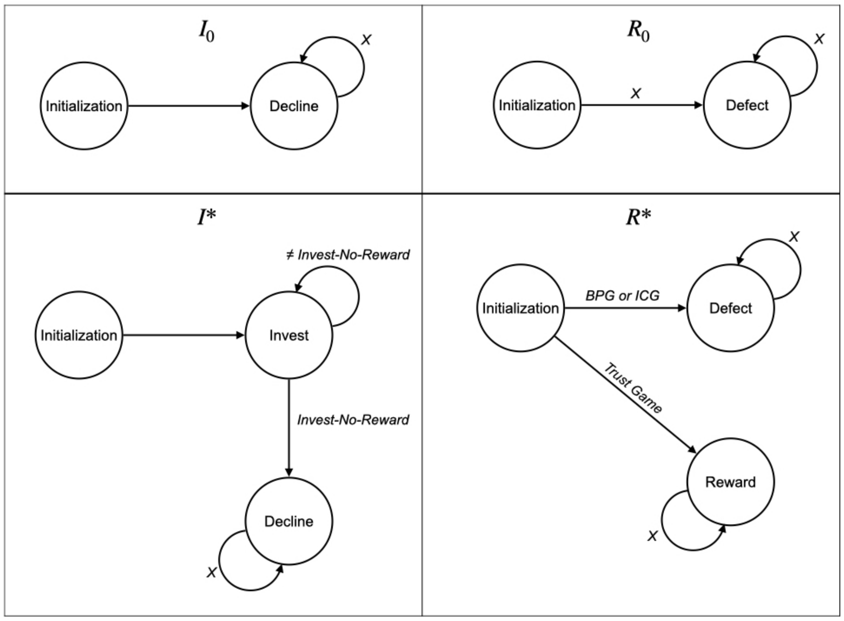

Figure 3: **Strategies in the investment game represented as finite automata**. Mutational distances are $d(I^*, I_0) = 5$, and $d(R^*, R_0) = 4$.

One can verify that, provided the probability that the game is an inefficient cooperation game is not too high, the profile (I^*, R^*) is a Nash equilibrium, and that, contrary to *Grim* or *TFT* in the prisoner's dilemma, if the probability that the game is a by-product game is high enough, (I^*, R^*) is parsimoniously evolvable from the pure defection profile (I_0, R_0). This is the case because the conditionality of the investor's behavior has an adaptive value in itself. Even if the responder is a pure defector, it is in the investor's interest to test an initial round of each game to see if the game is a by-product cooperation game.

Thus, the concept of adaptive parsimony does not imply that reciprocal cooperation is impossible. It implies that the ecological conditions under which reciprocal cooperation can evolve are more restrictive than in the prisoner's dilemma (see also André 2015; Bshary et al. 2016 and Ito et al. 2017).

Moreover, like the repeated prisoner's dilemma, this game has a large variety of subgame perfect equilibria in which the investor (i) invests in every round if he is in a by-product game, (ii) never invests if he is in an inefficient cooperation game, and (iii) invests in a variable fraction of rounds if he is in a trust game. One extreme equilibrium is where the investor invests in all rounds of the trust game and the responder systematically rewards him. Another extreme is where the investor never invests in the trust game and the responder never rewards him. And all possible intermediate equilibria exist between these two extremes, where the investor only invests in *some* rounds of the trust game and the responder only rewards these specific rounds and no others (see Appendix B).

Crucially, however, none of the intermediate equilibria are parsimonious. There are only two parsimonious equilibria. The first is fully cooperative: individuals cooperate in all rounds of the by-product game and extend this cooperation to all rounds of the trust game, which corresponds to the reciprocal cooperation profile (I^*, R^*) defined above. The second is completely uncooperative: individuals cooperate only in the by-product game and never extend this cooperation to the trust game, which corresponds to the complete absence of reciprocal cooperation. Intermediate equilibria are not parsimonious because cooperating in some rounds but not in others implies the ability to condition one's behavior on round number, which can never be gradually shaped by selection.

Finally, it is interesting to note that the parsimonious cooperative equilibrium profile (I^*, R^*) is not subgame perfect. As in all repeated games, subgame perfection is achieved when individuals are equipped with a self-punishing mechanism that ensures that sticking to the strategy after any deviation is individually rational. Such a self-punishing mechanism cannot be parsimonious because it also implies a conditional ability that cannot be gradually shaped by selection.

4 Discussion

In this article, I have argued that biological evolution cannot actually produce the wide variety of Nash equilibria predicted by game theory. The vast array of strategies that are both highly conditional and highly arbitrary may well be equilibria because they are self-stabilizing, but they cannot evolve through the gradual accumulation of adaptive mutations. I have proposed an equilibrium refinement, the concept of evolutionarily parsimonious equilibrium, that captures this idea. This refinement eliminates the vast majority of equilibria, especially in repeated games, and the equilibria it preserves are more biologically reasonable than others. I will now briefly discuss the relationship between this equilibrium refinement and other approaches to equilibrium selection that exist in the literature.

4.1 The cost of complexity

The concept of evolutionarily parsimonious equilibrium states that a complex conditional strategy cannot evolve if it has no added value over a simpler constitutive strategy. In this respect, it is closely related to game-theoretic approaches that consider the cost of complexity in the characterization of equilibria (Abreu and Rubinstein 1988; Banks and Sundaram 1990; Binmore and Samuelson 1992; Cooper 1996; Van Veelen and Garcıa 2019; Volij 2002).

The difference between the two approaches, however, is that considering a cost of complexity only eliminates conditional capabilities that are superfluous *in equilibrium*. In contrast, the concept of evolutionary parsimony also eliminates the possibility that superfluous conditional abilities play a role *in the path* to equilibrium. Even behavioral contingencies that end up being useful in equilibrium cannot be the result of evolution if there is no evolutionary pathway in which they have adaptive value when they first appear.

This distinction has fundamental implications. For example, in the repeated prisoner's dilemma, when noise is added, every strategy in the *Grim* family is an equilibrium according to the complexity cost criterion. On the contrary, even with noise, *Grim* is not evolutionarily parsimonious because it implies the emergence of a conditional ability without an advantage. Similarly, in the divide-the-dollar game (section 3.1) or the investment game (section 3.5), considering a cost of complexity does not reduce the diversity of equilibria, since all conditionalities, even arbitrary ones, are used in equilibrium. On the contrary, the parsimony approach eliminates most equilibria, since it stipulates that evolution cannot converge to arbitrary conditional capacities.

4.2 Subgame perfection

Subgame perfection requires that a strategy respond optimally to all possible histories, including histories that are not reached in equilibrium. This is related to the constraint imposed by the notion of parsimony that an equilibrium strategy must contain conditional capabilities that were adaptive on the path to equilibrium, i.e., under conditions other than those realized at equilibrium. For example, in the ultimatum game, the perfection condition and the parsimony condition have the same effect for the same reason. From the perfection point of view, rejecting an offer in equilibrium is suboptimal (even if it never happens). From the parsimony point of view: in the course of evolution, it can never be adaptive to reject offers, so evolution cannot converge on such a strategy. For this reason, in practice, the notion of perfection often leads to equilibria that are biologically reasonable because they are also parsimonious.

But the concepts of parsimony and perfection are also profoundly different. Perfection implies that a strategy responds optimally in all possible histories, assuming that the continuation game is played with a partner who plays the equilibrium strategy *itself*. In contrast, parsimony implies that the equilibrium strategy is adaptive in interactions with individuals playing *other* strategies encountered on the path to equilibrium, and it does not imply that a strategy responds optimally in histories that never occur in equilibrium. This has three consequences.

First, most perfect equilibria are not parsimonious. This is typically the case in repeated games. The *Grim* family in the repeated prisoner's dilemma, or the partial cooperation strategies in the investment game, are subgame perfect because they are optimal in all histories. But they are not parsimonious because they involve conditional capabilities that have no adaptive value on the path to equilibrium. This is why parsimony, but not perfection, dramatically reduces the diversity of equilibria in repeated games.

Second, parsimonious equilibria are often imperfect, since parsimony does not imply that a strategy responds optimally to game histories that never occur. This makes sense for a concept that seeks to reflect the consequences of biological evolution. Any action taken on a history that never occurs is neutral and has no reason to be optimized by selection. The concept of parsimony recognizes that evolution can lead to mismatches when organisms are placed in non-ecological situations.

Third, in repeated games, the criterion of parsimony is *incompatible* with perfection. Perfection is achieved thanks to a "trick" of strategies of the *Grim* type, which consists in self-punishment when one has deviated. Such a trick may well constitute an equilibrium when it is present, but it is never favored by evolution when it is initially absent. Therefore, in repeated games, evolution cannot converge to perfect strategies.

4.3 Stationary equilibria

Another equilibrium refinement considered in game theory and related to parsimony is the notion of stationarity. A stationary strategy is one in which players' actions depend only on the current state of the game and not on the history of past actions. Although it would require further analysis, stationarity is likely a consequence of parsimony. A stationary strategy does not condition its actions on arbitrary features of the past history of the game (such as having respected a complex pattern of actions in the past), but only on ecologically relevant properties of the environment, which is the essence of the idea of parsimony. For this reason, it is reasonable to expect that the concept of parsimony would also reduce the diversity of equilibria in the n-person alternating bargaining game. With $n > 2$, this game has a multiplicity of subgame perfect equilibria, but only one stationary equilibrium (Chatterjee et al. 1999), and probably only one evolutionarily parsimonious equilibrium, which would consist in dividing the resource into n equal shares.

4.4 Cultural group selection

As mentioned in the introduction, the multiplicity of equilibria plays an important role in evolutionary anthropology as a seemingly irrefutable argument that cultural group selection is necessary for human cooperation.

This has important implications for the nature of our social cognition. If reciprocity (and, more generally, repetition) can lead to any equilibrium, and does not particularly favor cooperative over noncooperative or partially cooperative equilibria, then human nature cannot include a universal, genetically encoded predisposition to cooperate, but only a predisposition to acquire whatever cultural norm–the local Nash equilibrium–is in the group in which one is born. In the words of N. Henrich and J. Henrich (2007): "*Some groups will develop norms about constructing community buildings, not eating snakes, and fishing in cooperative units, while other groups may culturally evolve norms about cooperatively raiding other groups, sending children to school, and giving young girls clitoridectomies*" (p. 68). The only mechanism that can favor cooperation is then ex post group selection, as Boyd and Richerson (2009) put it: "*Systems of reciprocity and reputation can stabilize a wide range of behaviors [...]. Rapid cultural adaptation can then lead to persistent differences between local social groups, [...] and then competition between groups leads to the spread of behaviors that increase the competitive ability of groups.*" Binmore (2009) also uses the same idea to explain the human sense of justice: "*Societies operating social contract a will therefore grow faster. [...] we will then eventually observe large numbers of copies of societies operating social contract a compared with those operating contract b*".

In this article, I have argued that this reasoning is incorrect because it dramatically exaggerates the problem of multiplicity of equilibria in biology. Given the constraint of evolutionary convergence, the vast majority of Nash equilibria are eliminated ex ante by the process of biological evolution, because they involve arbitrary rules of behavior that cannot be gradually shaped by selection. Thus, while the multiplicity of equilibria is a real problem in standard game theory, it does not translate into an equivalent problem in biology, which would require the intervention of nonparsimonious mechanisms such as group selection.

Acknowledgments. This study was supported by the EUR FrontCog grant ANR-17-EURE-0017 and the "Soutien à la mobilité internationale" of InSHS-CNRS.

References

Abreu, D. and A. Rubinstein (1988). "The Structure of Nash Equilibrium in Repeated Games with Finite Automata". In: *Econometrica* 56.6, p. 1259. ISSN: 00129682. DOI: 10.2307/1913097.

André (Mar. 2014). "Mechanistic constraints and the unlikely evolution of reciprocal cooperation". In: *Journal of Evolutionary Biology* 27.4. ISBN: 1010-061x, pp. 784–795. ISSN: 14209101. DOI: 10.1111/jeb.12351. URL: http://www.ncbi.nlm.nih.gov/pubmed/24618005 (visited on 03/19/2014).

— (2015). "Contingency in the evolutionary emergence of reciprocal cooperation". In: *The American Naturalist* 185.3, pp. 303–316. ISSN: 00030147. DOI: 10.1086/679625. URL: http://www.journals.uchicago.edu/doi/10.1086/679625 (visited on 02/11/2015).

André and T. Day (July 2007). “Perfect reciprocity is the only evolutionarily stable strategy in the continuous iterated prisoner’s dilemma”. In: *Journal of Theoretical Biology* 247.1, pp. 11–22. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2007.02.007. URL: http://www.ncbi.nlm.nih.gov/pubmed/17397874 (visited on 04/18/2012).

André and S. Nolfi (2016). “Evolutionary robotics simulations help explain why reciprocity is rare in nature”. In: *Scientific Reports* in press. Publisher: Nature Publishing Group, p. 32785. ISSN: 20452322. DOI: 10.1038/srep32785. URL: http://www.nature.com/articles/srep32785.

Aumann, R. J. (1974). “Subjectivity and correlation in randomized strategies”. In: *Journal of Mathematical Economics* 1, pp. 67–96.

Aumann, R. J. and L. S. Shapley (1994). “Long-Term Competition: A game-theoretic analysis”. In: *Essays in Game Theory in Honor of Michael Maschler*. Ed. by N. Megiddo. New York: Springer.

Aumann, R. J. (1998). “Correlated equilibrium as an expression of bayesian rationality”. In.

Banks and Sundaram (1990). “Repeated games, finite automata, and complexity”. In: *Games and Economic Behavior* 2.2, pp. 97–117. ISSN: 10902473. DOI: 10.1016/0899-8256(90)90024-O.

Bergmuller, R. et al. (2007). “Integrating cooperative breeding into theoretical concepts of cooperation”. In: *Behavioural Processes* 76.2. Publisher: Amsterdam: Elsevier Scientific Pub. Co., 1976-, pp. 61–72. DOI: 10.1016/j.beproc.2007.07.001.

Binmore (1985). “Bargaining and coalitions”. In: *Game-Theoretic Models of Bargaining*. Ed. by A. E. Roth. Cambridge: Cambridge University Press, pp. 269–304. ISBN: 978-0-521-26757-1. DOI: 10.1017/CBO9780511528309.014. URL: https://www.cambridge.org/core/books/gametheoretic-models-of-bargaining/bargaining-and-coalitions/6F6820B27297CD2B6AC021355F771DED (visited on 06/23/2023).

— (2009). “Fairness as a natural phenomenon”. In: *Philosophy* 8.22. Publisher: ESRC Centre for Economic Learning and Social Evolution, pp. 7–23. URL: http://eprints.ucl.ac.uk/14995/.

Binmore and Samuelson (Aug. 1992). “Evolutionary stability in repeated games played by finite automata”. In: *Journal of Economic Theory* 57.2, pp. 278–305. ISSN: 00220531. DOI: 10.1016/0022-0531(92)90037-I. URL: http://linkinghub.elsevier.com/retrieve/pii/002205319290037I.

— (1994). “An Economist’s Perspective on the Evolution of Norms”. en. In: *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft* 150.1, pp. 45–63. URL: http://www.jstor.org/stable/40753015.

Boyd, R. (2006). “Reciprocity: You have to think different”. In: *Journal of Evolutionary Biology* 19.5. ISBN: 1010061X, pp. 1380–1382. ISSN: 1010061X. DOI: 10.1111/j.1420-9101.2006.01159.x.

Boyd, R. and J. P. Lorberbaum (1987). “No Pure Strategy Is Evolutionarily Stable in the Repeated Prisoners-Dilemma Game”. In: *Nature* 327.6117, pp. 58–59.

Boyd, R. and P. J. P. J. Richerson (Nov. 2009). “Culture and the evolution of human cooperation”. In: *Philosophical Transactions of the Royal Society B-Biological Sciences* 364.1533, pp. 3281–3288. ISSN: 1471-2970. DOI: 10.1098/rstb.2009.0134. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2781880&tool=pmcentrez&rendertype=abstract (visited on 05/26/2013).

Bshary, R. et al. (2016). “Why mutual helping in most natural systems is neither conflict-free nor based on maximal conflict”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 371.1687. ISSN: 14712970. DOI: 10.1098/rstb.2015.0091.

Chatterjee, K. et al. (1999). “N-Person Bargaining and Strategic Complexity”. In: April 1998, pp. 1–45.

Connor, R. C. (1995). “The benefits of mutualism: a conceptual framework”. In: *Biological Reviews* 70.3. Publisher: Wiley Online Library, pp. 427–457.

Cooper, D. J. (Jan. 1996). “Supergames Played by Finite Automata with Finite Costs of Complexity in an Evolutionary Setting”. en. In: *Journal of Economic Theory* 68.1, pp. 266–275. ISSN: 00220531. DOI: 10.1006/jeth.1996.0015. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022053196900150 (visited on 06/27/2023).

Dawkins, R. (1996). *Climbing Mount Improbable*. Norton.

Eshel, I. (1983). “Evolutionary and continuous stability”. In: *Journal of Theoretical Biology* 103, pp. 99–111.

Eshel, I. et al. (Apr. 1997). “Continuous stability and evolutionary convergence.” In: *Journal of theoretical biology* 185.3, pp. 333–43. ISSN: 0022-5193. DOI: 10.1006/jtbi.1996.0312. URL: http://www.ncbi.nlm.nih.gov/pubmed/9156083.

Fisher, R. A. (1930). “The Genetical Theory of Natural Selection”. In: Oxford: Clarendon Press, pp. 73–83.

Fudenberg, D. and E. Maskin (1990). “Evolution and Cooperation in Noisy Repeated Games”. In: *American Economic Review* 80.2. ISBN: 00028282, pp. 274–279. ISSN: 00028282. URL: http://www.jstor.org/stable/2006583.

Gale, J. et al. (Jan. 1995). “Learning to be imperfect: The ultimatum game”. en. In: *Games and Economic Behavior* 8.1, pp. 56–90. ISSN: 08998256. DOI: 10.1016/S0899-8256(05)80017-X. URL: https://linkinghub.elsevier.com/retrieve/pii/S089982560580017X (visited on 06/23/2023).

García, J. and A. Traulsen (Apr. 2012). “The structure of mutations and the evolution of cooperation”. In: *PloS one* 0.1. Ed. by A. Szolnoki, pp. 1–4. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0035287. URL: http://dx.plos.org/10.1371/journal.pone.0035287 (visited on 06/11/2013).

García, J. and M. Van Veelen (Jan. 2016). “In and out of equilibrium I: Evolution of strategies in repeated games with discounting”. en. In: *Journal of Economic Theory* 161, pp. 161–189. ISSN: 00220531. DOI: 10.1016/j.jet.2015.11.007. URL: https://linkinghub.elsevier.com/retrieve/pii/S002205311500201X (visited on 06/27/2023).

Geoffroy, F. and J. B. André (2021). "The emergence of cooperation by evolutionary generalization". In: *Proceedings of the Royal Society B: Biological Sciences* 288.1951, pp. 1–9. ISSN: 14712954. DOI: 10.1098/rspb.2021.0338.

Geritz, S. A. et al. (1998). "Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree". In: *Evolutionary Ecology* 12.1, pp. 35–57. ISSN: 02697653. DOI: 10.1023/A:1006554906681.

Gintis, H. (Feb. 2007). "A framework for the unification of the behavioral sciences." In: *The Behavioral and brain sciences* 30.1, 1–16, discussion 16–61. ISSN: 0140-525X. DOI: 10.1017/S0140525X07000581. URL: http://www.ncbi.nlm.nih.gov/pubmed/17475022.

Henrich, N. and J. Henrich (2007). *Why Humans Cooperate.* arXiv: 1011.1669v3 Series Title: Evolution and Cognition ISSN: 0717-6163. Oxford: Oxford University Press. ISBN: 978-0-19-530068-0.

Ito, K. et al. (2017). "The evolution of cooperation by negotiation in a noisy world". In: *Journal of Evolutionary Biology.* ISSN: 14209101. DOI: 10.1111/jeb.13030.

Lehmann, L. and L. Keller (2006). "The evolution of cooperation and altruism – a general framework and a classification of models". In: *Journal of Evolutionary Biology* 19, pp. 1365–1376.

Maynard Smith, J. and G. A. Parker (1976). "Logic of Asymmetric Contests". In: *Animal Behaviour* 24.FEB, pp. 159–175.

Nash, J. F. (1950). "The Bargaining Problem". In: DOI: 10.4324/9780429303005-7.

Noë, R. and P. Hammerstein (1994). "Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating". In: *Behavioral Ecology and Sociobiology* 35.1, pp. 1–11. ISSN: 03405443. DOI: 10.1007/BF00167053.

Nowak, M. A., A. Sasaki, et al. (2004). "Emergence of cooperation and evolutionary stability in finite populations". In: *Nature* 428.6983, pp. 646–650.

Nowak, M. A. and K. Sigmund (1992). "Tit for tat in heterogeneous populations". In: *Nature* 355.6357, pp. 250–253. ISSN: 00280836. DOI: 10.1038/355250a0.

Riley, J. G. (2001). "Silver signals: Twenty-five years of screening and signaling". In: *Journal of Economic Literature* 39.2, pp. 432–478. ISSN: 00220515. DOI: 10.1257/jel.39.2.432.

Roberts, G. et al. (2021). "The benefits of being seen to help others: indirect reciprocity and reputation-based partner choice". In: *Philosophical Transactions of the Royal Society B.* ISBN: 0000000219433.

Rubinstein, A. (Jan. 1982). "Perfect Equilibrium in a Bargaining Model". en. In: *Econometrica* 50.1, p. 97. ISSN: 00129682. DOI: 10.2307/1912531. URL: https://www.jstor.org/stable/1912531?origin=crossref (visited on 06/23/2023).

Selten, R. (1965). "Spieltheoretische Behandlung eines Otigopolmodells mit Nachfragetrfigheit". In: *Zeitschrift f{ü}r Gesamte Staatswissenschaft* 121, pp. 301–324.

Selten, R. (1973). “A Simple Model of Imperfect Competition, where 4 are Few and 6 are Many”. In: *International Journal of Game Theory* 2, pp. 141–201.

— (1975). “Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games”. In: *International Journal of Game Theory* 4, pp. 25–55.

— (1983). “Evolutionary stability in extensive two-person games”. In: *Mathematical Social Sciences* 5, pp. 269–363.

— (1988). “Evolutionary stability in extensive two-person games - correction and further development”. In: *Mathematical Social Sciences* 16, pp. 223–266.

Shaked, A. (1986). *A Three-Person Unanimity Game, talk given at the Los Angeles national meetings of the Institute of Management Sciences and the Operations Research Society of America.*

Skyrms, B. (1996). *Evolution of the social contract*. Publication Title: Evolution of the Social Contract. Cambridge: Cambridge University Press. ISBN: 978-1-139-92482-5. DOI: 10.1017/CBO9781139924825.

Sobel, J. (2012). “Signaling games”. In: *Computational Complexity: Theory, Techniques, and Applications* 9781461418. ISBN: 9781461418009, pp. 2830–2844. DOI: 10.1007/978-1-4614-1800-9_174.

Van Veelen, M. (2012). “Robustness against indirect invasions”. In: *Games and Economic Behavior* 74.1. Publisher: Elsevier Inc., pp. 382–393. ISSN: 08998256. DOI: 10.1016/j.geb.2011.05.010. URL: http://dx.doi.org/10.1016/j.geb.2011.05.010.

Van Veelen, M. and J. Garcıa (2019). “In and out of equilibrium II: evolution in repeated games with discounting and complexity costs.” en. In: *Games and Economic Behavior* 115, pp. 113–130.

Volij, O. (May 2002). “In Defense of DEFECT”. en. In: *Games and Economic Behavior* 39.2, pp. 309–321. ISSN: 08998256. DOI: 10.1006/game.2001.0893. URL: https://linkinghub.elsevier.com/retrieve/pii/S0899825601908930 (visited on 06/27/2023).

Appendix

A Measuring mutational distance with finite automata

Here I seek a formal approach to obtaining a proxy for the mutational distance between two strategies. No approach is perfect, since the true mutational distance is a biochemical variable that cannot be measured from a simple model.

To this aim, I first characterize each strategy as a finite-state machine (see Figures 1, 2, 3, 4, and 5). For the sake of conceptual clarity (and without losing generality), I always consider an initial state before the start of the game, called the initialization state, which is present in all games and all strategies.

To measure the distance between two strategies, I then have to solve an alignment problem. To do this, I try all possible alignments between the two strategies and define the distance between them as the smallest distance found in all these trials, i.e. with the best possible alignment. More precisely, I proceed as follows:

Let there be two strategies S_1 and S_2 with k_1 and k_2 states respectively (with $k_2 k_1$), and let there be an alignment $\mathcal{A}$ defined as a bijective mapping from the set of states of S_1 to the image of that set in the states of S_2 (i.e, each state of S_1 has one and only one image in S_2), subject only to the constraint that the image of the initialization state of S_1 must be the initialization state of S_2. Any pair consisting of a state of S_1 and its image in S_2 according to $\mathcal{A}$ is called a pair of homologous states according to $\mathcal{A}$. States in S_2 which are not the image of any state in S_1 are said to be non-homologous according to $\mathcal{A}$.

I then measure the following 4 quantities with this alignment:

- Δ_B: Number of pairs of homologous states where the behavior of S_2 is different from that of S_1.
- Δ_E: Number of non-homologous states of S_2 $(k_2 - k_1)$.
- Δ_N: Total number of transitions from non-homologous states of S_2.
- Δ_T: Number of differences between the two strategies among transitions from homologous states.

The distance between S_1 and S_2 according to $\mathcal{A}$ is given by $d_{\mathcal{A}}(S_1, S_2) = \Delta_B + \Delta_E \Delta_N + \Delta_T$. The distance between S_1 and S_2 is then obtained by minimizing $d_{\mathcal{A}}$ over all possible alignments, i.e., $d(S_1, S_2) = \min\limits_{\mathcal{A}} d_{\mathcal{A}}(S_1, S_2)$. Figures 1, **??** and 3 show examples of this approach.

In principle, another possible approach to measuring the distance between two strategies might be to consider all possible information states in the game from the start, not just those that each strategy really distinguishes, and to count the number of such states in which the two strategies play different actions.

For example, in the repeated prisoner's dilemma with memory 1, the individuals can be in two information states: one's partner has just cooperated or one's partner has just defected. In this case, pure defection is defined by the pair of actions (D, D) since it plays D in each of these states, pure cooperation is defined by the pair (C, C) and tit-for-tat is defined as an intermediate strategy that plays C only in one of the two states and D in the other, i.e., by the pair (C, D). By this measure of distance, full defection would be considered closer to tit-for-tat than to full cooperation.

This way of measuring distances is implicitly used in many models of evolutionary game theory (e.g., Lehmann and Keller 2006; Nowak and Sigmund 1992), and it has been shown to have important implications for evolutionary dynamics (André 2014; García and Traulsen 2012). While this may be considered a reasonable way to *describe* strategies, I contend that it is not reasonable from a mechanistic point of view. From a machine that constitutively performs a single action, creating a machine that constitutively performs another action requires fewer modifications than creating a machine that can first measure a property of the environment and then respond to that property by conditionally performing two possible actions. This insight is captured by the measure of distance I propose above. By this measure, full defection is indeed closer to full cooperation than to tit-for-tat (see Figure 2) because full cooperation simply plays a different action than full defection in its single state, whereas tit-for-tat has both 1 additional state and 3 modified or supplementary transitions.

B Investment game

The investment game is an asymmetric repeated game played by two individuals, called the Investor and the Responder. The game is preceded by a move by Nature that randomly determines the type of stage game that the two individuals will play repeatedly (see main text, section 3.5). Three types of stage games are possible: the stage game can be a by-product cooperation game, a trust game or an inefficient cooperation game, with respectively probabilities p_b, p_t and p_i, all strictly positive.

Regardless of the game type chosen by Nature, the investor has two possible actions in each round: *Invest* or *Decline*. Then, if the investor chooses to invest in a given round, the responder has two possible actions: *Reward* or *Defect*. If the investor chooses to decline, then the responder has no decision to make.

In each round of the game, several events are possible: (1) "Invest-No-Reward", when the investor has invested and received no reward, neither from his partner nor from Nature. (2) "Invest-Reward", when the investor has invested and received a reward, either from his partner or from Nature, this event being itself the union of two events: (3) "Invest-Natural-Reward" and (4) "Invest-Social-Reward". (5) "Decline", when the investor has not invested in this round.

The payoffs for each game are as follows:

In the trust game, the cost of investing is c, the benefit of receiving the investment is b, the cost of rewarding is r, and the benefit of receiving the reward is also r (i.e., the reward is a conservative transfer), and we assume that we have the relation $b > r > c > 0$.

In the by-product game, the automatic benefit of investing is $a > 0$, and all other parameters are as in the trust game.

In the inefficient cooperation game, for simplicity and without loss of generality, I assume that the benefit of receiving the investment is 0, and that all other parameters are as in the trust game.

Finally, for simplicity and without loss of generality, I assume that the game is repeated for an expected number of rounds n, which is infinite. Thus, any strictly positive benefit accrued over the whole game will always be greater than any finite cost paid only once. This assumption could easily be relaxed. n would then simply have to be greater than some finite computable threshold. However, this would make the presentation unnecessarily cumbersome.

B.1 Investors cannot distinguish the two types of rewards

The investor can receive two types of rewards in this game: a social reward, which he receives when he is in a trust game if the responder decides to reward him, and a natural reward, which he receives automatically when he is in a by-product cooperation game. In this first version of the model, I assume for simplicity that, for some exogenous reason beyond his control, it is impossible for the investor to distinguish between the two types of reward. Apart from this restriction, both players have access to all information about past actions.

Each strategy profile consists of a pair of strategies, the investor's strategy and the responder's strategy. Consider the ancestral strategy profile (I_0, R_0), where the investor unconditionally plays *Decline* and the responder unconditionally plays *Defect.* This strategy is represented by a finite automaton in Figure 3.

Consider then the strategy profile $S^* = (I^*, R^*)$ (see Figure 3), defined as follows:

1. Investor's strategy I^*:
   - If the "Invest-No-reward" event never happened in the past, play *Invest*
   - In all other situations play *Decline*
2. Responder's strategy R^*:
   - If the first round of the game was a by-product cooperation game or an inefficient cooperation game, play *Defect.*
   - If the first round of the game was a trust game, play *Reward*

B.1.1 S^* is Nash equilibrium

On the investor's side, if the investor invests in the first round, he can discover the nature of the game he is playing, so in the long run he will win $n[p_t(r-c)+p_b a]-p_i c$, whereas if he does not invest, he gets 0, which is always lower under our assumptions. Thus, deviating from I^* by not investing in the first round is costly. In subsequent rounds, if there have been no "invest-no-reward" events, then the game is a by-product cooperation game or a trust game, and in this case, in every round, the investment payoff is strictly positive in expectation, so there is no benefit to deviating from I^*. On the other hand, if there has been at least one "invest-no-reward" event in the past, then the game is an innefficient cooperation game, and in this case the expected payoff of investing is negative, so there is no point in deviating from I^*.

On the responder's side, if the first round of the game is a by-product cooperation game or an innefficient cooperation game, then every round will be a by-product cooperation game or an innefficient cooperation game, since Nature doesn't change its mind. In this case, there is no profit in rewarding the investor, either because the investor will invest in every round anyway (if the game is a by-product cooperation game), or because his investment will not yield a net profit (if the game is an innefficient cooperation game). Therefore, there is no profit to be made by deviating from R^*. Conversely, if the first round is a trust game, rewarding the investor leads to a payoff $n(b-r) > 0$, while deviating leads to a zero payoff.

In conclusion, no deviation is strictly favored on either side. The profile (I^*, R^*) is a Nash equilibrium.

B.1.2 S^* is an evolutionarily parsimonious equilibrium

To show that the strategy profile (I^*, R^*) is parsimoniously evolvable from (I_0, R_0), I will first show that the profile (I^*, R_0) is parsimoniously evolvable from (I_0, R_0) in one step, and then show that the profile (I^*, R^*) is parsimoniously evolvable from (I^*, R_0) in one step.

To prove that (I^*, R_0) is parsimoniously evolvable from (I_0, R_0) in one step, one must show that there is no alternative strategy I' such that $d(I', I_0) < d(I_0, I^*)$ and $P(I', R_0) \geq P(I^*, R_0)$, where $P(I^*, R_0) = np_b a - (p_t + p_i)c$, which is strictly greater than zero under our assumption.

For a strategy to be closer to I_0 than I^*, it must necessarily give up its conditionality. So there are only two possibilities. Either I' plays *Decline* unconditionally in all rounds, i.e. $I' = I_0$. In this case $P(I', R_0) = 0$, which by hypothesis is less than $P(I^*, R_0)$. Or I' plays *Invest* unconditionally in all rounds. Its distance from I_0 is then $d(I', I_0) = 1$, which is smaller than $d(I^*, I_0) = 5$. But I' gets a payoff $P(I', R_0) = n[-(p_t + p_i)c + p_b a]$, which is lower than $P(I^*, R_0)$, because I' repeatedly invests in games where it gets a negative payoff.

Tthere is no alternative strategy I' such that $d(I', I_0) < d(I_0, I^*)$ and $P(I', R_0) \geq P(I^*, R_0)$. The profile (I^*, R_0) is thus parsimoniously evolvable from (I_0, R_0) in one step.

To prove that (I^*, R^*) is parsimoniously evolvable from (I^*, R_0) in one step, one must show that there is no alternative strategy R' such that $d(R', R_0) < d(R^*, R_0)$ and $P(R', I^*) \geq P(R^*, I^*)$, where $P(R^*, I^*) = n[p_b b + p_t(b - r)]$, with $b > r > c$ and $n(b - r) > b$.

Again, there are only two possibilities. Either R' plays *Defect* unconditionally in all rounds, i.e. $R' = R_0$. In this case, $P(R', I^*) = (np_b + p_t)b$. Since we have $n(b - r) > b$, this payoff is strictly lower than $P(R^*, I^*)$, i.e., it is always worthwhile for a responder to reward in the trust game in order to receive long-term investments, rather than taking advantage of a single round. Or R' plays *Reward* unconditionally in all rounds. In this case we have the distance $d(R', R_0) = 1$, which is smaller than $d(R^*, R_0) = 4$, but we have the payoff $P(R', I^*) = n[(p_b + p_t)(b - r) - p_i r]$, which is also strictly less than $P(R^*, I^*)$ as long as $r > 0$ and $p_i + p_b > 0$ (which is the case by assumption).

There is no alternative strategy R' such that $d(R', R_0) < d(R^*, R_0)$ and $P(R', I^*) \geq P(R^*, I^*)$. The profile (I^*, R^*) is thus parsimoniously evolvable from (I^*, R_0) in one step.

In conclusion, the strategy profile (I^*, R^*) is parsimoniously evolvable (in two steps) from (I_0, R_0). Since (I^*, R^*) is also a Nash equilibrium, it is an evolutionarily parsimonious equilibrium from (I_0, R_0).

B.1.3 S^* is not Subgame perfect

It is easy to see that (I^*, R^*) is not a subgame perfect strategy profile. In the out-of-equilibrium situation where a responder deviates from R^* and plays *Defect* once in a trust game, the strategy I^* dictates that the investor should then play *Decline* in all future rounds of the game, even though the best response would be to play *Invest*, since the responder will return to the R^* profile and reward all future investments.

A more complex strategy on the responder's side, R^{**}, would be necessary to obtain a strategy profile (I^*, R^{**}) that is subgame perfect. To achieve this, R^{**} must contain a self-punishing principle as follows:

- If the first round of the game was a by-product cooperation game or an inefficient cooperation game, play *Defect.*
- If the first round of the game was a trust game and the responder has always rewarded in the past, play *Reward.*
- If the first round of the game was a trust game, but the responder has defected at least once in the past, play *Defect*

However, since the behavior of R^{**} depends on a larger number of parameters than R^*, the mutational distance between R^{**} and R_0 is larger than between R^* and R_0. In the framework of finite automata, we get the distance $d(R^{**}, R_0) = 6$, whereas we had $d(R^*, R_0) = 4$ (see Figure 4).

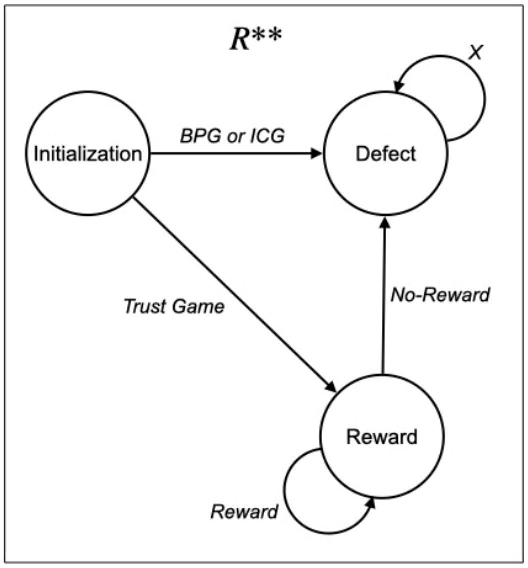

Figure 4: **Subgame perfect strategy on the responder's side in the investment game**

So we have $P(R^*, I^*) = P(R^{**}, I^*)$, but $d(R^*, R_0) < d(R^{**}, R_0)$, i.e. compared to R^{**}, R^* is both closer to R_0 and just as good in terms of payoff. The profile (I^*, R^{**}) is therefore not parsimoniously evolvable from (I^*, R_0) in one step.

More generally, the responder's ability to punish himself gives him strictly no advantage under any circumstances. It is at best neutral. This does not prevent this strategy profile from being subgame perfect, but it does prevent it from being an evolutionarily parsimonious equilibrium (and thus a plausible product of evolution), since there are no circumstances in which this ability is favored by directional selection.

B.2 Investors can distinguish the two types of rewards

In this second version of the model, the investor can distinguish two types of rewards: the natural rewards that he automatically receives in the by-product cooperation game, and the social rewards that he receives from trustworthy partners in the trust game.

Consider a subset $A \subset Z^+$ of strictly positive integers, and consider the strategy profile (I^*_A, R^*_A) defined as follows (see Figure 5 for a representation as finite automata):

1. Investor's strategy I^*_A:
    - Play *Invest* if at least one of the following three conditions is true:
        - This is the first round of the game.
        - There is at least one round in the past where the "Invest-Natural-Reward" event occurred
        - The current round number belongs to A, and there are no past rounds that belong to $\{A \cup 1\}$ in which the event "Invest-No-Reward" occurred.
    - Play *Decline* in all other situations
2. Responder's strategy R^*_A:
    - Play *Reward* if the current round number belongs to A and the first stage game was a trust game.
    - Play *Defect* in all other situations

Note that with this definition, if the subset A is the set of all natural numbers, then the strategy profile becomes $(I^*_{Z^+}, R^*_{Z^+}) = (I^*, R^*)$.

B.2.1 For any infinite size $A \subset Z^+$, the profile (I^*_A, R^*_A) is a Nash equilibrium

Consider an infinite size $A \subset Z^+$. Since I assume that the game is repeated an infinite number of times n, it is also repeated an infinite number of times in rounds belonging to A. And, whatever a round number k, since A is infinite, there are an infinite number of rounds after k belonging to A. The game is therefore repeated an infinite number of times in A after k.

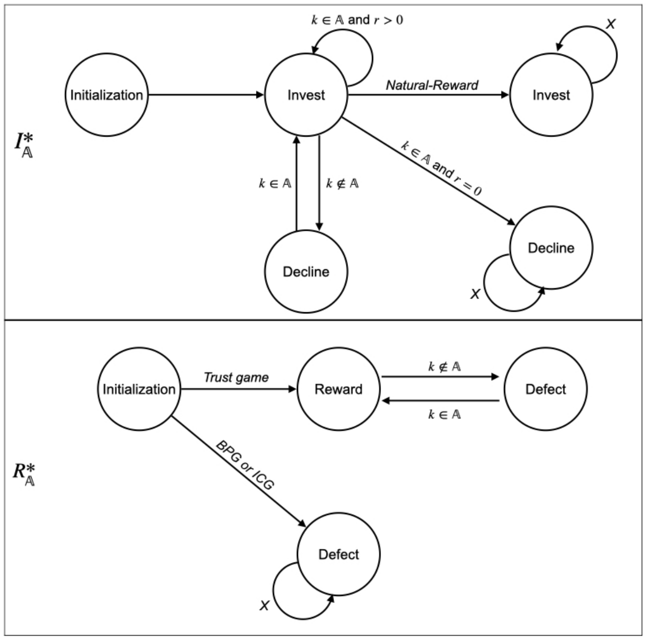

Figure 5: **Partial cooperation strategies in the investment game.**

Here I aim to show that (I^*_A, R^*_A) is a Nash equilibrium. To this end, I consider all possible deviations on both sides and verify that they cannot yield a strictly positive benefit.

On the investor's side:

(1) In the first round of a game, I^*_A dictates to play *Invest*. If the investor plays *Invest*, he can discover the type of game he is playing, and so in the long run he will make a strictly positive profit as long as the game is repeated for long enough, which we have assumed. Therefore, deviating from I^*_A is costly.

(2) If there is at least one round in the past in which the event "Invest-Natural-Reward" occurred, I^*_A dictates to play *Invest*. If this situation is reached, it implies that every stage game is a by-product game, because Nature does not change its mind. Thus, in every stage game, investing yields payoff $a > 0$, so there is no advantage in deviating from I^*_A.

(3) If the current round number belongs to A, and there are no past rounds belonging to A in which the "Invest-No-Reward" event occurred, I^*_A dictates to play *Invest*. If this situation is reached, it implies that either the partner or Nature will reward every investment, so there is no point in deviating from I^*_A.

(4) If the current round number does not belong to A and the investor has never received a natural reward in the past, I^*_A dictates to play *Decline*. If this situation is reached, it implies that neither Nature nor the partner will reward this investment. So there is no point to deviate from I^*_A.

(5) If the current round number belongs to A, but there is at least one past round

belonging to A in which the event "Invest-No-Reward" occurred, I_A^* dictates to play *Decline*. If this situation is reached, it implies that the stage game is an inefficient cooperation game, in which case the partner will not reward the investment. So there is no point in deviating from I_A^*.

On the responder's side:

(1) If the first stage game was a by-product cooperation game or an inefficient cooperation game, R_A^* dictates to play *Defect* in all rounds. It cannot be beneficial to deviate because rewarding has no benefit in these games.

(2) If the first stage game was a trust game, each stage game will be a trust game. The partner will invest in all rounds belonging to A and only in those rounds, and R_A^* dictates to play *Reward* after each investment. Deviating from R_A^* by defecting for one round would give a one-shot gain, but cause a long-run cost by ending cooperation. Since we have assumed that the game will be repeated long enough for a profit accumulated over the whole game to always be greater than a cost paid only once, it cannot be beneficial to deviate from R_A^*.

In conclusion, no deviation is favored. The profile (I_A^*, R_A^*) is a Nash equilibrium for any infinite size $A \subset Z^+$.

B.2.2 The profile $(I_\emptyset^*, R_\emptyset^*)$ is a Nash equilibrium

On the investor's side:

In the first round, if the investor plays *Invest* as stipulated by $I_\emptyset^*$, he can discover the type of game he is playing, and so in the long term he will get a strictly positive payoff as long as the game is repeated for long enough, which we have assumed. Deviating from $I_\emptyset^*$ would thus be costly.

In subsequent rounds, if there is at least one round in the past in which the event "Invest-Natural-Reward" occurred, $I_\emptyset^*$ stipulates to invest. Indeed, this situation implies that every stage game is a by-product game. Thus, in every stage, investing yields payoff $a > 0$, so there is no gain in declining instead. On the other hand, if the investor has never received a natural reward, $I_\emptyset^*$ stipulates to decline. And indeed, this situation implies that neither Nature nor the partner will ever reward. So there is no benefit to invest.

On the responder's side: Regardless of the nature of the stage game, the investor's behavior is completely independent of the responder's actions, so there is never any benefit to reward.

In conclusion, no deviation is favored. The profile $(I_\emptyset^*, R_\emptyset^*)$ is a Nash equilibrium.

B.2.3 The profile $(I_\emptyset^*, R_\emptyset^*)$ is an evolutionarily parsimonious equilibrium

First, note that $R_\emptyset^* = R_0$. This is simply a responder that never actively rewards.

So to show that $(I_\emptyset^*, R_\emptyset^*)$ is a parsimonious equilibrium from (I_0, R_0), we only need to show that $(I_\emptyset^*, R_0)$ is parsimoniously evolvable from (I_0, R_0) in one step.

We have $P(I^*_\emptyset, R_0) = P(I^*, R_0)$, i.e., against the ancestral strategy R_0, $I^*_\emptyset$ is strictly equivalent to I^* in terms of payoff. However, we also have $d(I^*_\emptyset, I_0) = d(I^*, I_0)$, i.e., $I^*_\emptyset$ is no further away from I_0 than I^* was. Since (I^*, R_0) is parsimoniously evolvable from (I_0, R_0) in one step, then $(I^*_\emptyset, R_0)$ is also parsimoniously evolvable from (I_0, R_0) in one step.

So $(I^*_\emptyset, R^*_\emptyset)$ is parsimoniously evolvable from (I_0, R_0). And since it's a Nash equilibrium, it's a parsimonious equilibrium.

B.2.4 $\forall A \notin \{\emptyset, Z^+\}$, the profile (I^*_A, R^*_A) is not parsimoniously evolvable from (I_0, R_0)

We know that $d(I^*_A, I_0) > d(I^*, I_0)$ (see Figure **??**), but we have $P(I^*_A, R_0) = P(I^*, R_0)$. In other words, an investor who behaves conditionally on the round number is (i) further away from the ancestor in terms of mutational distance but (ii) does not get a larger payoff than an investor who invests regardless of round number. Hence (I^*_A, R^*_A) is not parsimoniously evolvable from (I_0, R_0) in one step.

More generally, a strategy that conditions its action on the fact that the round number belongs to A can never be parsimoniously evolvable in one step from an ancestor that does not, since this supplementary conditional ability confers no advantage. Thus (I^*_A, R^*_A) could not even be parsimoniously evolvable from (I_0, R_0) in multiple steps. There is no evolutionary path from (I_0, R_0) to (I^*_A, R^*_A) by a sequence of parsimonious adaptive mutations.

B.2.5 Multiplicity of subgame perfect equilibria

As in the simpler version of the game, none of the parsimonious strategies is subgame perfect. But we can easily construct a subgame perfect strategy profile. Consider again an infinitely large set $A \subset Z^+$, and the strategy profile (I^*_A, R^{**}_A), where

I^*_A is defined as above, and R^{**}_A is defined as follows: Play *Reward* iff the stage game is a trust game and one of the following two conditions is true:

- This is the first round of the game.
- The current round number belongs to A, and the responder has always played *Reward* in all rounds belonging to $\{A \cup 1\}$ in the past

In all other cases, play *Defect*

This strategy profile is subgame perfect for any set $A \subset Z^+$, since no deviation is strictly advantageous, even on out-of-equilibrium histories. That is, any level of cooperation can be supported by a subgame perfect equilibrium, from the complete absence of reciprocal cooperation (i.e., players cooperate only in the by-product game) to the maximum possible level of reciprocal cooperation (players cooperate in all rounds of both the by-product and the trust game), including all intermediaries (players cooperate only in a fraction of rounds of the trust game). None of these equilibria is parsimonious, however, because they include a self-punishment mechanism that cannot be gradually shaped by selection.

library