bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Individual selection leads to collective efficiency through coordination

Arthur Bernard[a,c], Nicolas Bredeche[a], and Jean-Baptiste André[b]

September 21, 2018

a. Sorbonne Université, CNRS, Institut des Systèmes Intelligents et de Robotique, F-75005 Paris, France

b. Institut Jean Nicod, Département dâĂŹétudes cognitives, ENS, EHESS, PSL Research University, CNRS, Paris France

c. Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland

Correspondence: jeanbaptisteandre@gmail.com and nicolas.bredeche@sorbonne-universite.fr

Keywords: Evolution of cooperation | Equilibrium selection | Coordination game | Collective action | Evolutionary robotics | Evolutionary game theory | Indirect genetic effects

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Evolutionary game theory shows that social interactions involving coordination between individuals are subject to an "evolutionary trap." Once a suboptimal strategy has evolved, mutants playing an alternative strategy are counterselected because they fail to coordinate with the majority. This creates a detrimental situation from which evolution cannot escape. To determine how this problem materializes in a model with a greater degree of realism than conventional game-theoretical models, we simulate the life and the long-term evolution of a population of individuals playing a two-player coordination game, using the framework of evolutionary robotics. We first confirm the existence of an evolutionary trap in a simple setting. We then, however, reveal that this problem disappears in a more realistic setting where individuals need to coordinate with one another. In this setting, robots evolve an ability to adapt plastically their behavior to one another, as this improves the efficiency of their interaction. This ability has an unintended evolutionary consequence: a genetic mutation affecting one individual's behavior also indirectly alters their partner's behavior because the two individuals influence one another, which is called an "indirect genetic effect". Consequently, pairs of partners can virtually change strategy together with a single mutation, and the evolutionary barrier between alternative strategies disappears. This finding reveals a general principle that could play a role in nature to smoothen the transition to efficient collective behaviors in all games with multiple equilibriums.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 Introduction

The success of a collective action often hinges on the coordinated decisions of several individuals. For instance, carrying out a collective hunt implies that all individuals hunt at the same time, agree on a common prey, and pursue the prey in a coordinated manner. Thus, collective efficiency does not depend solely on the skills of a single individual but emerges from the ability of the group to act together (Alvard and Nolin, 2002; Alvard, 2003; Drea and Carter, 2009). This begs the question of how natural selection, which acts on *individuals*, can shape such *collective* behaviors.

This problem can be formalized with a specific class of games called "coordination games" To understand, let us consider a situation in which two hunters must coordinate to capture a prey, but have to make a choice between a prey with a high nutritive value and a prey with a low nutritive value. The strategy of choosing the most nutritious prey is evolutionarily stable. If everyone chooses this prey, one's best response is to choose this prey as well. But choosing the poorly nutritious prey is also evolutionarily stable. If everyone chooses the poorly nutritious prey, there is nothing better one can do than choose the same. The existence of this second, suboptimal, evolutionarily stable strategy (ESS) raises a problem because evolution can hardly move from one ESS to another. If all hunters initially target the low-value prey, mutants preferring the high-value prey are counterselected by frequency-dependent selection because they fail to coordinate with the majority. Hence, individuals are trapped in a suboptimal ESS and collective efficiency is not maximized.

All collective actions where individuals need to coordinate with one another, and where they can do so in either an efficient or an inefficient way, entail such an "evolutionary trap." Inefficient coordinated behaviors evolve that cannot later be improved by natural selection because *individual* selection has no way of improving *collective* efficiency in a coordination game.

Evolutionary game theoreticians and evolutionary biologists have explored two hypotheses that can explain

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

how this problem can be solved in nature. The first hypothesis is based on stochastic effects (Foster and Young, 1990; Kandori et al., 1993; Young, 1993; see also Nowak et al., 2004 in a different setting). In a finite population fixed in a particular ESS, counterselected mutants can rise in frequency due to genetic drift, and eventually destabilize the existing ESS, thereby moving the population away from the evolutionary trap, toward another, generally superior, ESS. The second hypothesis is based on group selection (Boyd and Richerson, 1990, 2002; Richerson and Boyd, 2005; Boyd and Richerson, 2009). Due to chance, different groups of individuals may initially evolve different ESSes of the same game. If these groups compete with one another, the groups that happen to play the most efficient ESS will eventually prevail, allowing this strategy to spread in the entire population. In sum, according to available theories, coordination games suffer from an evolutionary trap problem, and collective efficiency in these games can ensue either from demographic stochasticity or from group selection, but not from plain individual selection.

However, so far, coordination games have been formally studied in models that were highly stylized, in particular with regard to the mechanistic underpinning of behavior, and these simplifications may have important consequences. In this paper, we describe simulations of a coordination game using evolutionary robotics (Nolfi and Floreano, 2000; Doncieux et al., 2015). As compared to classic evolutionary game-theoretical approaches, evolutionary robotics provides a more realistic modeling of individuals and their environment (Mitri et al., 2013; Trianni, 2014), capturing in particular the practical problems raised by coordination (Bernard et al., 2016).

In this setting, we show that the evolutionary trap actually disappears altogether. Robots evolve a behavioral solution to coordinate with one another that generates indirect genetic effects (Wolf et al., 1998), hereby changing "the rules of the game". Through their own behavior, robots transform a coordination game with multiple ESSes and an evolutionary trap problem into a simple individual optimization problem with a single, maximally efficient, ESS. In this transformed game, collective efficiency is reached by plain individual selection, with no need for genetic drift or group selection. We posit that many collective optimization problems may be solved

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in similar ways in nature.

2 Results

We simulate a collective hunt in which two players must coordinate and attack together the same prey to gain a benefit. The two-dimensional environment contains two types of prey, poorly nutritious prey called *boars* (worth 125 payoff units for each individual hunter), and highly nutritious prey called *stags* (worth 250 units for each hunter). Hunting alone is possible, but it provides 0 payoff unit (cf. payoff matrix in Table 1). This game features two Nash equilibriums (and thus two ESSes): to hunt either boars or stags, with the latter equilibrium providing a higher payoff. In technical terms, hunting stags is called the "payoff-dominant" equilibrium.

The robots we use as players are each driven by a multilayer Perceptron Rumelhart et al. (1986) that maps sensory inputs to motor outputs, with neural weights subject to artificial evolution. Each robot is endowed with proximity sensors all around its body. These sensors are capable, within a limited range, of discriminating between boars, stags, the other robot, and walls. So as to maintain the prey density constant, a captured boar (or stag) is removed from the environment and relocated to a new position (*Methods*).

2.1 Individual selection cannot generate collective efficiency, in a simple setting

The first question we address is whether the evolutionary transition from the least efficient to the most efficient ESS can occur.

First, we let 30 independent populations of robots pre-evolve for 3000 generations with only 1 boar and 1 stag, and modified payoff values: hunting stags temporarily yields no reward. We thus ensure that these 30 populations all evolve the boar-hunting equilibrium, with all individuals always targeting boars and avoiding

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

stags.

Second, each of these 30 populations of evolved boar hunters is used as the seed for another 6000 generations of evolution, with the regular rewards for each prey reinstated (Table 1). In spite of its collective superiority, stag hunting never evolves within the next 6000 generations for any of 30 independent replicates (Figure 1). In every replicate, the mean proportion of stags hunted remains at 0 throughout the 6000 generations. Hence, individuals are genuinely trapped in the suboptimal equilibrium. That is, collective efficiency cannot ensue from plain individual selection.

2.2 Collective efficiency can be achieved by individual selection, in a more complex setting

In practice, predators are unlikely to live in a world with a single prey of each kind. In a realistic environment, hunters must agree on a specific *individual* prey to hunt (Alvard and Nolin, 2002; Alvard, 2003; Drea and Carter, 2009), and not just on the *type* of prey.

To investigate the consequences of this complication, we follow the same procedure as before. We pre-evolve 30 independent populations of pure boar hunters, using modified payoff values, with the stags never bringing any reward, but this time in an environment with several (9) identical boars and several (9) identical stags present.

We then let each of these 30 populations evolve for another 6000 generations with regular payoff values (Table 1) in the same environment with several boars and stags present (*Methods*).

In this setting, we observe that the transition from boar hunting to stag hunting does occur in 12 replicates out of 30 (Figure 2). This significantly differs from the previous results obtained in a simpler environment (Mann-Whitney U test on the number of replicates where the transition happened p-value <0.0001). Environ-

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

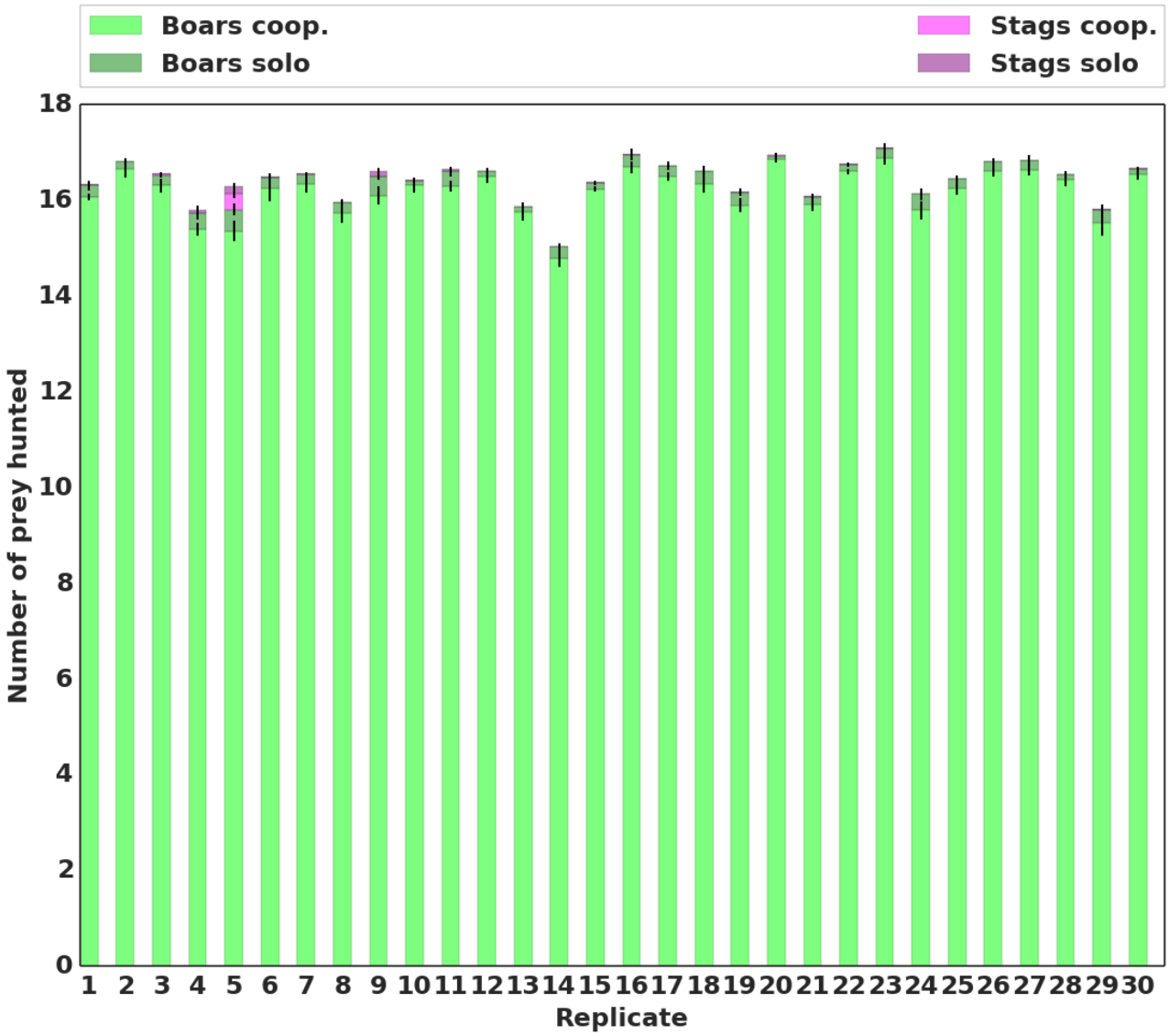

Figure 1: **Mean proportion of prey hunted.** Repartition of the prey hunted by the best individual in each replicate, at the last generation of evolution (generation 6000). We differentiated between the type of prey hunted (boar or stag). Rewards were 125 for a boar and 250 for a stag (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mental complexity promotes the evolutionary transition toward the pay-off dominant equilibrium in 40% of the replicates.

Taking a closer look at these 12 "successful" replicates reveals a particular kind of coordination strategy for collective hunting. Because the environment is more complex, individuals need to react to each other's behavior to stay together and converge on the same prey. To this end they evolve a behavioral strategy, which we refer to as the "turning" strategy, whereby they constantly turn around one another. This strategy ensures that they keep their partner in their line of sight and move toward a prey at the same time. Due to their proximity, an individual who gets on a prey is likely to be joined quickly by their partner (Figure 3, a video of this strategy is also available in Supplementary Materials).

This behavioral strategy has an evolutionary implication. Because hunters react to each other's behavior, a mutation affecting the behavior of one individual can also modify the behavior of her partner. A mutant attracted to stags rather than boars may thus succeed in hunting by transforming, albeit temporarily, her partner into a stag hunter as well. As a result, the evolutionary transition away from the suboptimal trap is facilitated in comparison with the simple environment.

2.3 Division of labor further facilitates collective optimization

The turning strategy obtained so far results from both hunters demonstrating identical behaviors, whether related to moving together or selecting a prey. While this similarity in behavior makes it possible for hunters to coordinate toward hunting the same prey, both individuals spend a significant amount of time turning around one another. This results in a tedious process of targeting a particular prey. The question is open as to the existence of more efficient coordination patterns, in particular with respect to assuming complementary behavioral strategies.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

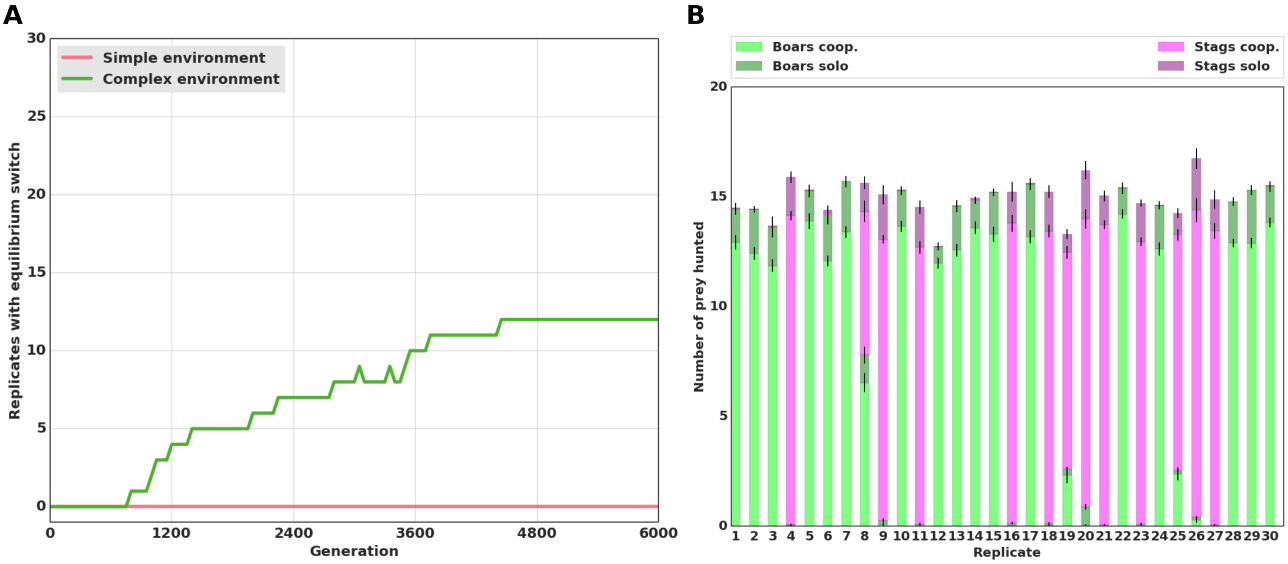

Figure 2: **Number of stag-hunting replicates and proportion of prey hunted.** *(A)* Number of replicates (out of a total of 30) where stag hunting evolved in the *Simple environment* and *Complex environment* settings. We considered that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the Simple environment setting, the environment was constituted of one boar and one stag. In comparison, in the Complex environment setting, 18 prey were present in the environment, and it was thus necessary to coordinate for cooperation to be possible. *(B)* Repartition of the type of prey hunted by the best individual in every replicate, at the last generation of evolution (generation 6000), in the Complex environment setting. Rewards were 125 for a boar and 250 for a stag (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

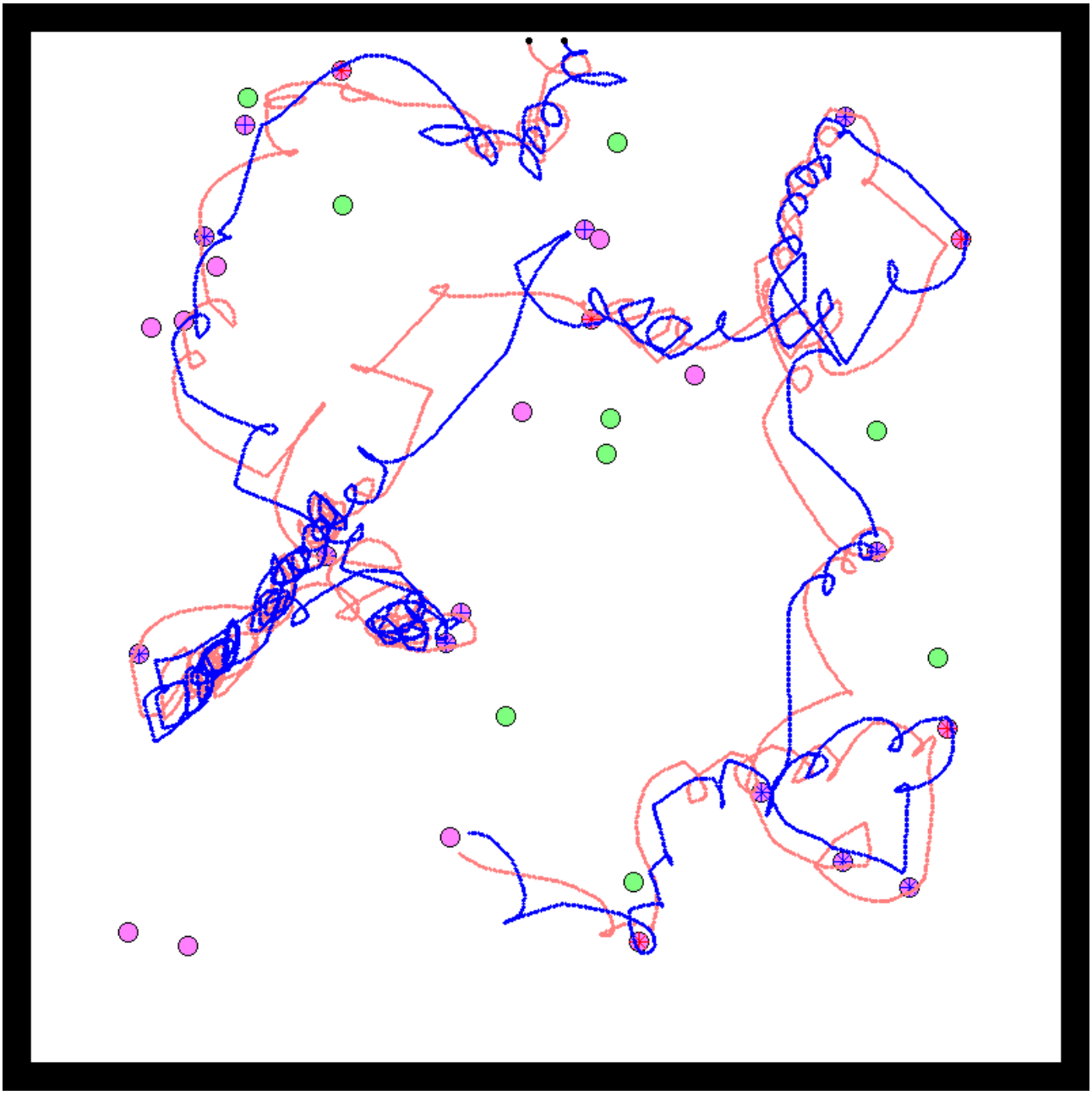

Figure 3: **Example of the turning strategy.** Example of a simulation where both individuals adopted a *turning* strategy. The path of the agents during this simulation is represented in red and blue, starting from their initial positions (represented by black dots). Each disc represents a prey in the environment. Boars are represented in green and stags in purple. When a prey was killed cooperatively, a red cross (resp. blue) is shown on the prey if the red agent (resp. blue) arrived at this prey first.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We posit that one possible limitation is the lack of expressivity of our choice of control function, a limitation that may not exist in nature. Though multilayered Perceptrons are theoretically universal approximators, this is hardly the case in practice (Cybenko, 1989). In particular, the ability to switch from one behavioral pattern to a completely different one may be required for breaking behavioral symmetry between two individuals, but may be hindered by the limitation of the controller used thus far.

To explore the possible benefits of more complex decision-making capabilities, we enable each robot with the possibility to choose from two (possibly very) different controllers, depending on the context at hand. To do so, we introduce a new evolutionary operator: the *network duplication* operator. Loosely inspired by gene duplication (Hughes, 2002), a newly created individual may be subject to the complete duplication of its artificial neural network. Conversely, an individual with two networks may have one deleted. In this setup, any individual may possess either one or two network(s). Whenever two individuals who possess two networks each interact with one another, we ensure that each of them expresses a different copy of their own networks (*Methods*).

We use the same experimental procedure as before: 30 populations are pre-evolved independently where only boar hunting yields a reward, and this time we introduce network duplications. We observe a significant difference with respect to previous experiments, as in *all* 30 replicates, we observe the evolution of asymmetrical hunting behaviors in the form of a *leader-follower* division of labor.

Results are also significantly different than in both previous treatments when the original payoff matrix is reinstated (Table 1), because the transition to stag hunting occurs in 22 replicates out of 30 (Figure 4) as compared to 12 replicates without network duplication. As in the pre-evolution runs, the *leader* guides the pair toward a given prey, always arriving first, while the *follower* keeps the leader in its line of sight at all times and joins her afterward on the prey (Figure 5, a video of this strategy is also available in Supplementary Materials).

This cognitive division of labor stems directly from the duplication of the neural controller: once duplicated,

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

one version of the network always ends up encoding for the leader behavior, while the other encodes for the follower behavior, just like duplicated genes of the same family encode slightly different functions. However, network duplication alone (i.e., without asymmetrical behavior) is not enough, as demonstrated by an additional control experiment where network duplication is allowed, but with both robots using the same network, randomly chosen at the start of each new evaluation (see Figure 7 and *Methods*).

The division of labor has two consequences. First, it improves hunting efficiency. In the turning strategy, the symmetry of decision-making sometimes hinders the ability to reach a consensus. Even though turning promotes coordination, individuals often still fail to converge on the same prey. In comparison, performance is significantly higher in the leader-follower strategy (Figure 6, Mann-Whitney U test on the mean reward at last generation, p-value <0.001): the frequency of coordination failures is reduced thanks to a clear separation of roles.

Second, the division of labor has an evolutionary consequence. In the leader-follower strategy, just as in the turning strategy, hunters react to each other's behavior and are therefore also prone to react to mutants' behavior. But, in contrast to the turning strategy, this response is asymmetrical and, therefore, more precise. Any mutation affecting the leader's behavior also changes completely the behavior of the follower. That is, a mutation in a single individual automatically affects two individuals at the same time. As a consequence, the adaptive valley between boar hunting and stag hunting disappears or, put differently, boar hunting ceases to be an equilibrium. Any increase in the probability of hunting a stag rather than a boar, when playing the role of a leader, is directly favored by individual selection, and pure stag hunting therefore becomes the only evolutionary equilibrium.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

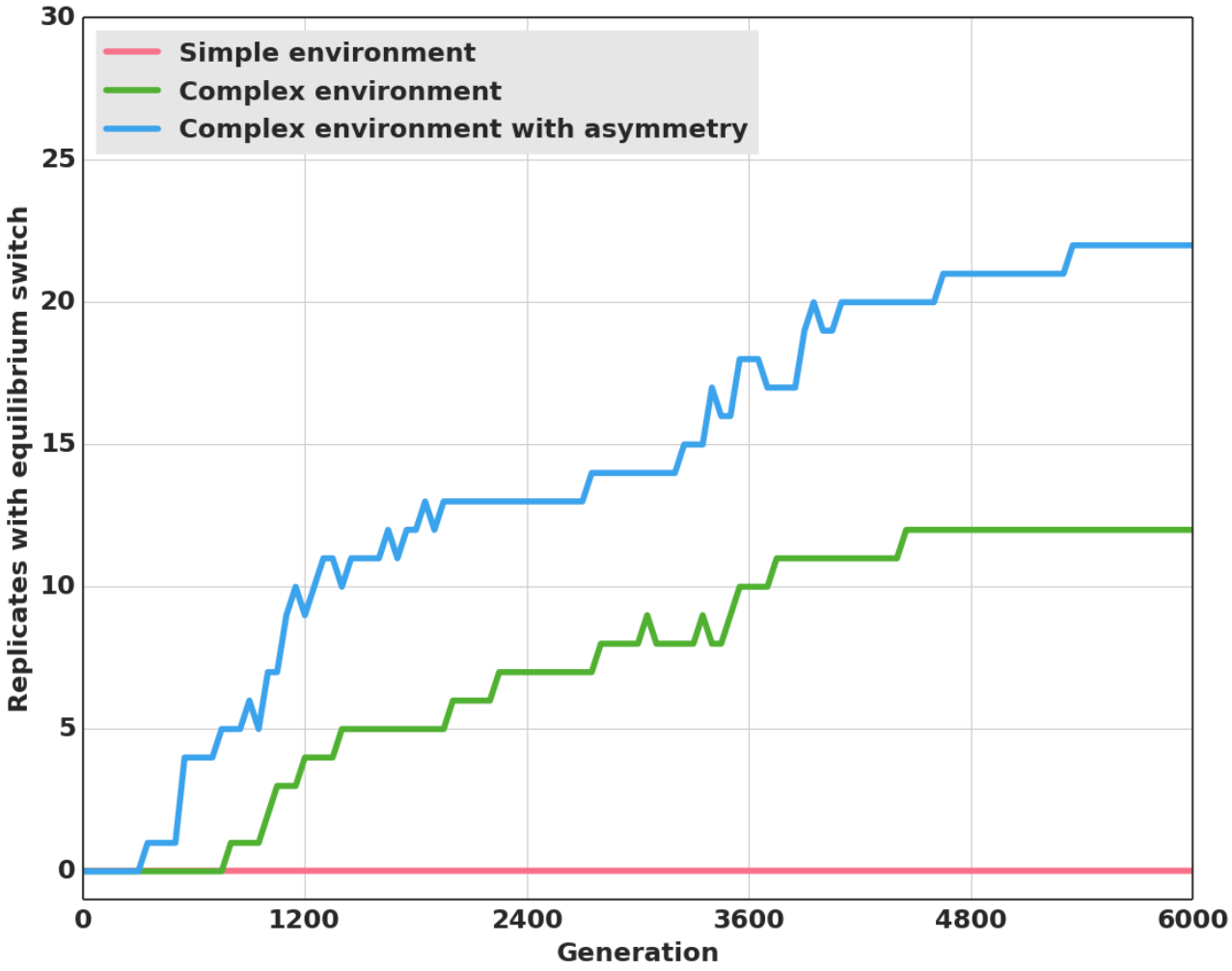

Figure 4: **Transitions to the optimal equilibrium through evolutionary time.** Number of replicates (out of a total of 30) where the switch to stag hunting occurred in the *Simple environment*, *Complex environment* and *Complex environment with asymmetry* settings. We considered that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. The Simple environment and Complex environment settings are the same as previously (Figure 2). In the Complex environment with asymmetry, a duplication event could occur where individuals would coevolve two neural networks. We enforced the asymmetry by forcing each hunter to use a different random neural network as its controller. Rewards were 125 for a boar and 250 for a stag (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

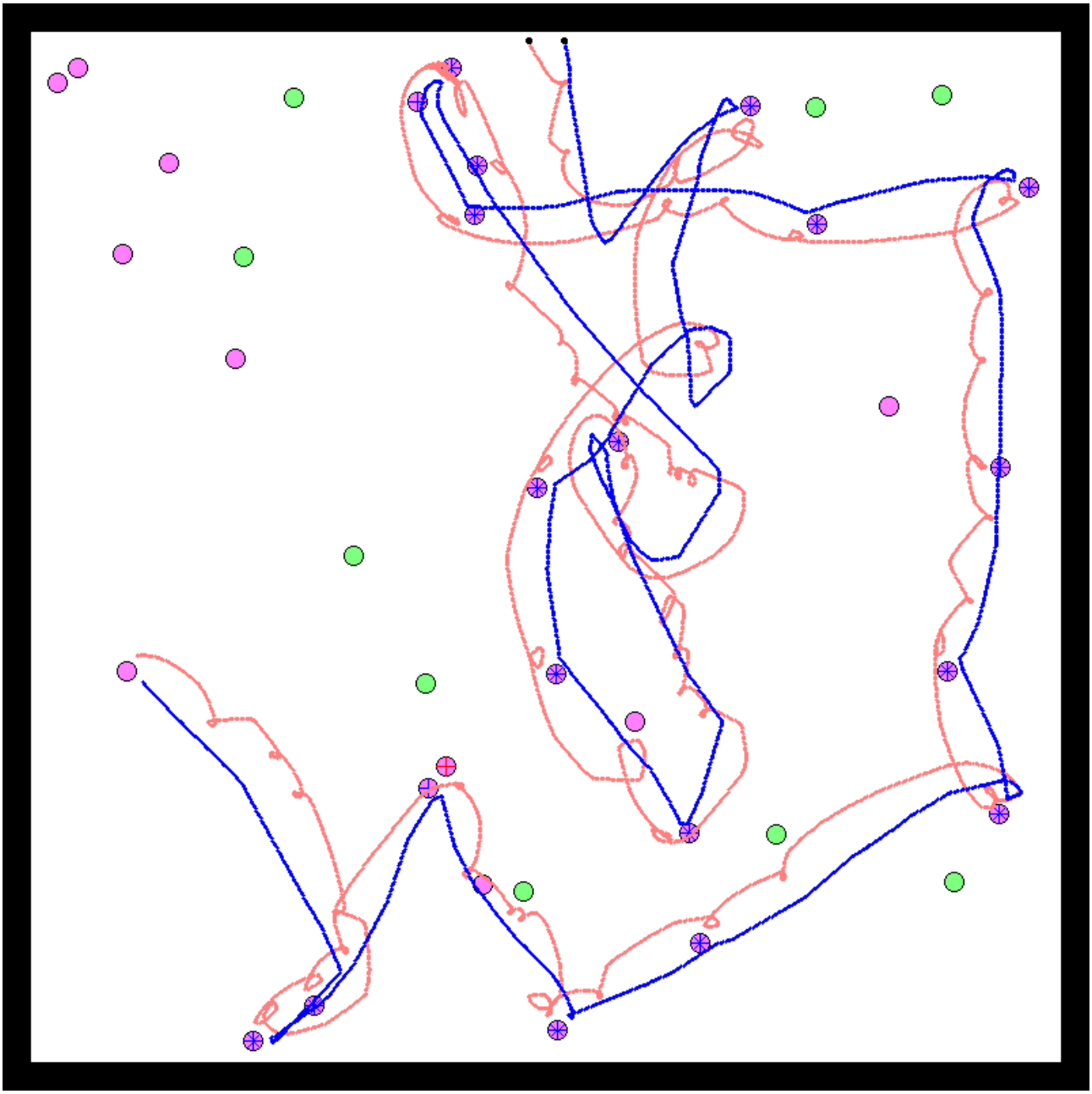

Figure 5: **Example of the leader-follower strategy.** Example of a simulation where both individuals adopted a *leader-follower* strategy. The path of the agents during this simulation is represented in red and blue, starting from their initial positions (represented by black dots). Each disc represents a prey in the environment. Boars are represented in green and stags in purple. When a prey was killed cooperatively, a red cross (resp. blue) is shown on the prey if the red agent (resp. blue) arrived on this prey first.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

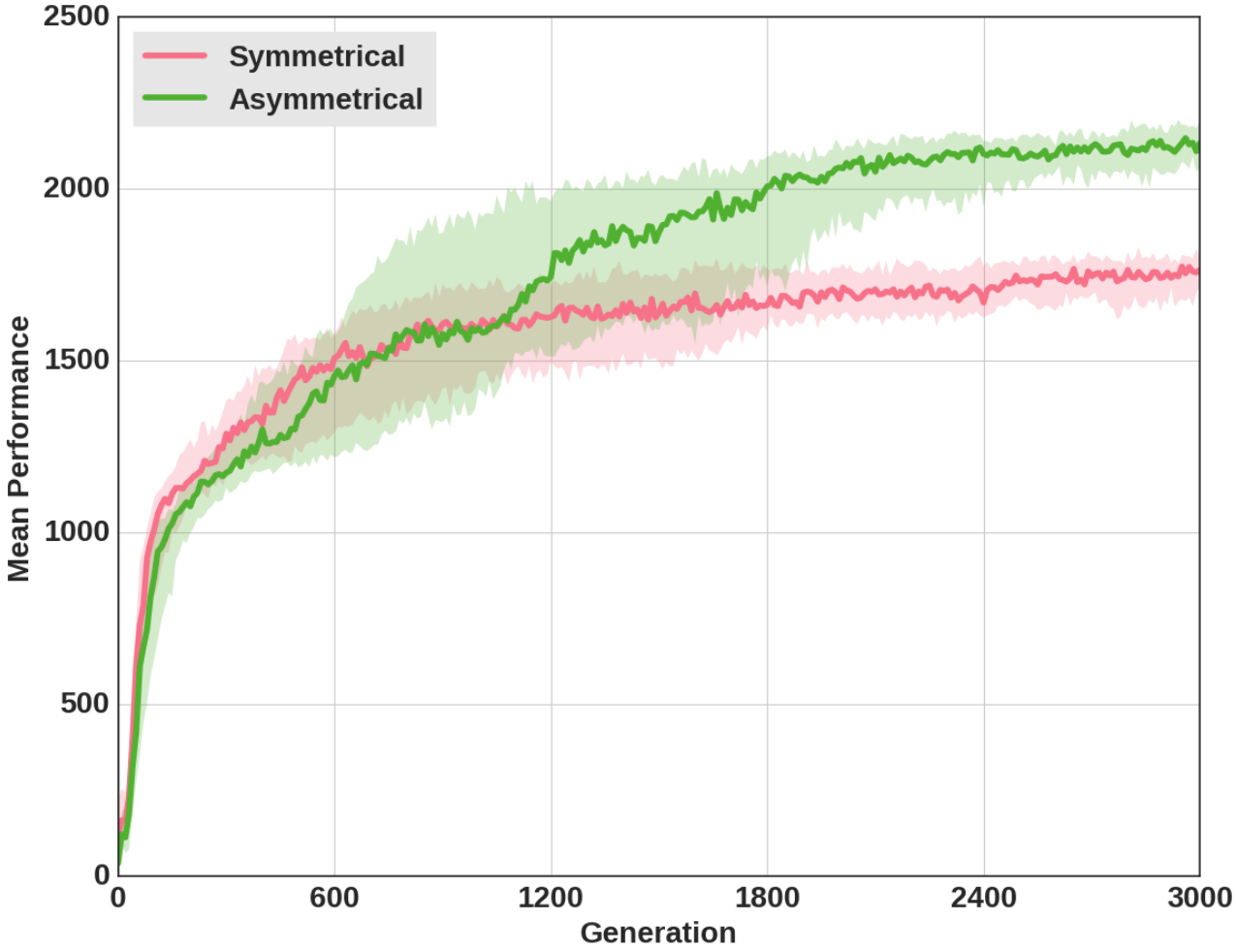

Figure 6: **Mean reward comparison of the turning strategy and the leader-follower strategy.** Mean reward over evolutionary time of the best individuals in each of the 30 replicates when adopting a turning strategy or a leader-follower strategy during the pre-evolution step (which lasted for 3000 generations). Rewards were 125 for a boar and 250 for a stag (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

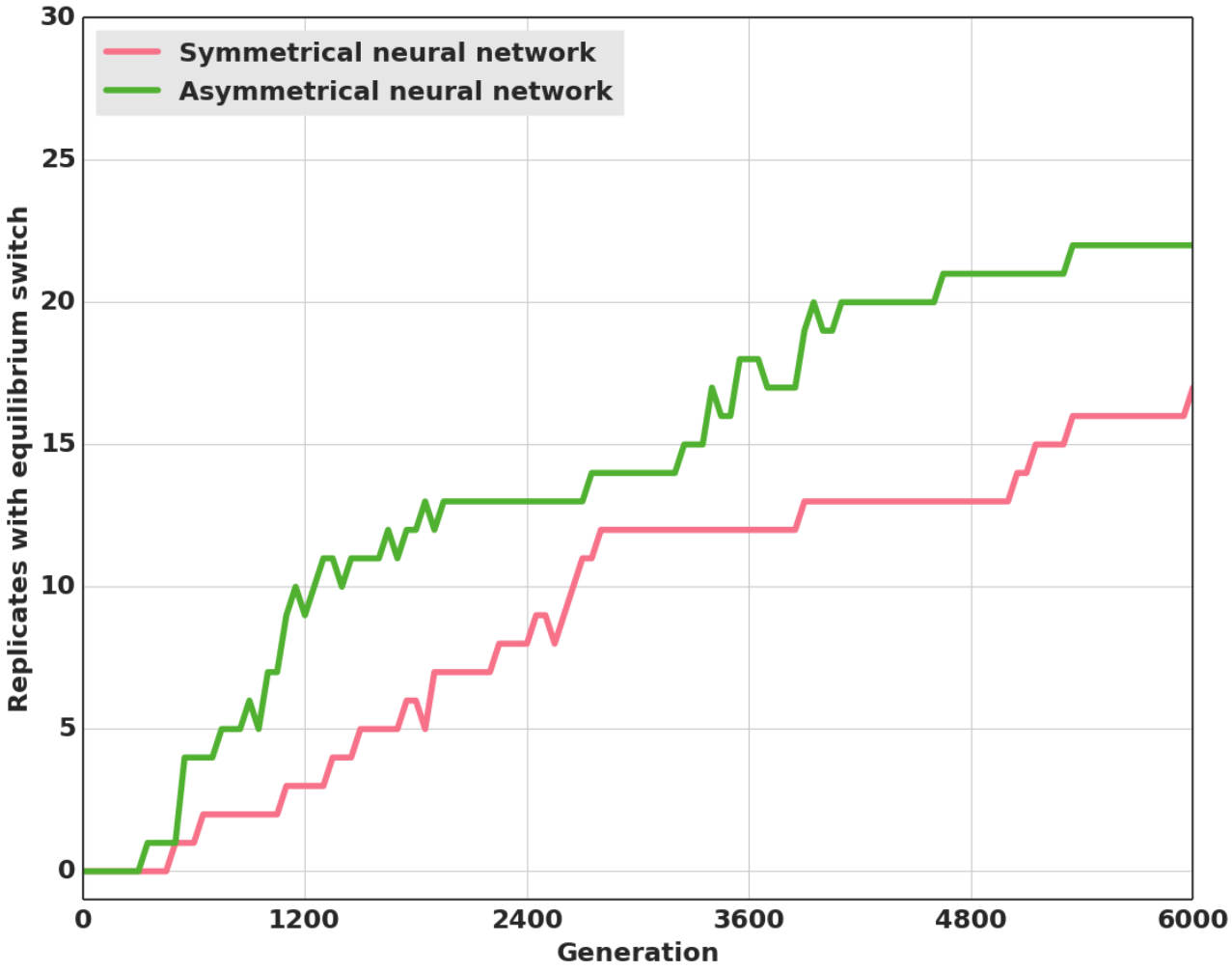

Figure 7: **Transitions to the optimal equilibrium through evolutionary time.** Number of replicates (out of a total of 30) where the switch to stag hunting occurred with a *symmetrical* or *asymmetrical* neural network. We considered that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In both treatments, a duplication event could occur where individuals would coevolve two neural networks. In the symmetrical treatment, both individuals were forced to use the same randomly chosen neural network during a simulation. In the asymmetrical treatment, individuals employed a different neural network for the entirety of the simulation. Rewards were 125 for a boar and 250 for a stag (Table 1).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3 Discussion

Collective actions often require several individuals to make coordinated choices. As a result, their efficiency, or lack of efficiency, is a collective property, not a property of any particular individual. This raises an evolutionary difficulty because natural selection acts on individual, not collective, properties. Collective actions are thus subject to an "evolutionary trap" problem. Once a relatively successful but still perfectible collective organization has evolved, any single mutant playing a better strategy will be counterselected due to her lack of coordination with others. For collective efficiency to be reached by evolution, several individuals would all somehow have to "mutate collectively", but genetic mutations do not occur in several organisms at the same time.

In this paper, we studied this problem in artificial robotics experiments. We simulated the life and the long-term evolution of a population of simple robots that played a 2 x 2 coordination game. Robots were hunters who could choose between two types of prey that were either poorly nutritious or highly nutritious. But they could only be successful if they converged together on the same prey. Hence, they faced a coordination problem with two ESSes-hunting poorly nutritious prey or hunting highly nutritious prey, and an adaptive valley in between. Our aim was to find out how the evolutionary trap problem materializes and how it is solved-or not solved-in a model possessing a greater degree of realism than conventional models of coordination games.

We first confirmed the existence of an evolutionary trap. In a simple setting where the environment was constituted of two individual prey only-one poorly nutritious and one highly nutritious-if we initially forced the robots to play the suboptimal ESS (attacking the poorly nutritious prey), all populations of robots remained stuck in this ESS "forever," that is, at least for the 6,000 generations of our simulations. Individual mutants who were targeting the better prey could not be favored due to their singularity.

This observation may seem at odds with evolutionary game-theoretical "drift" models (Foster and Young, 1990; Kandori et al., 1993; Young, 1993). According to these models, finite populations always eventually escape

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from evolutionary traps, because counterselected mutants rise in frequency by genetic drift and eventually replace the suboptimal resident. Mathematical analyses of this process show that, in the long run, populations should spend most of the time in the vicinity of one specific equilibrium (called the "stochastically stable" equilibrium), which corresponds to stag hunting with our parameter setting. Hence, "drift" models predict that our populations should not be trapped in the suboptimal strategy.

This discrepancy is a matter of time scale, however. According to drift models, our robots should eventually escape from the evolutionary trap and hunt stags, in the "long" run, but the question is how "long" in practice? The answer to that question depends a great deal on the practical availability of mutants. In game-theoretical models, stag hunters are simply assumed to occur by mutation from boar hunters at a given rate. In a robotic setting like this one, however, stag hunters must appear by random changes in the connection weights of boar hunters' neural networks, and multiple such changes separate a pure and well-optimized boar hunter from a pure and well-optimized stag hunter. As a result, mutants playing the stag-hunt strategy are extremely rare in a population of pure boar hunters.

This is confirmed in a supplementary experiment (see Fig. 8, and *Methods*), where we analyzed the behavior of 10^5 mutants generated randomly from a pure and well-optimized "boar hunter" genotype. This analysis showed that, at best, the random mutants merely had a *probabilistic* tendency to target the stag. That is, they played a *mixed* rather than a pure strategy. These intermediate mutants can never prevail in a population of pure boar hunters, even after a phase of genetic drift, because mixed strategies are strongly counterselected in coordination games (owing to the uncertainty they generate). Hence, these mutants cannot bring about the stochastic transition to stag hunting.

Although the occurrence of strong-effect mutants able to destabilize the suboptimal equilibrium is possible in principle, the above analysis shows that it is highly unlikely in practice owing to the mutational distance

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

between ESSes. As a result, the stochastic evasion from the evolutionary trap is an extraordinarily slow process in our simulations. Because the mutational distance between ESSes will often be even greater in biology than in the present setting, stochastic evasions from suboptimal equilibriums are presumably a highly improbable event in general in biological settings.

However, we then showed that this problem actually disappears in a more realistic setting. In a richer environment constituted of several prey of each kind-several poorly nutritious prey and several highly nutritious prey, individuals needed to actively coordinate with their partner to converge on the same prey. To resolve this problem, they evolved behavioral tactics to keep track of and follow their partner. In our experiments we observed the evolution of two such tactics. In the first series of experiments, individuals constantly turned around one another, never moving away from their partner, which increased the probability that they both would eventually converge on a prey. In other experiments where we authorized a behavioral asymmetry between partners, individuals evolved a leader/follower strategy whereby a single individual chose a prey, whereas the other simply followed her.

These coordination strategies evolved because they had immediate individual benefits. They increased the probability for individuals to hunt successfully. But they also had an unintended evolutionary consequence. When individuals had the capacity to coordinate with each other, a mutation affecting the behavior of one individual also indirectly modified, phenotypically, the behavior of his/her partner, almost as if individuals had mutated "collectively." In quantitative genetics, such an effect is called an "indirect genetic effect" (Wolf et al., 1998) because a gene affects the phenotype of an individual in which it is not directly expressed. Indirect genetic effects are well known for changing the evolutionary process in sometimes dramatic ways, by altering the genotype-phenotype relationship. In the present case, behavioral coordination tactics evolved to deal with the uncertainty of the "normal" environment in which one's partners only targeted suboptimal prey but the precise individual prey they were targeting could vary, which required being able to follow them. However, once evolved,

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

coordination tactics also happened to work efficiently when interacting with mutants who preferentially targeted other types of prey. They led one to follow and coordinate with mutants like they did with "normal" residents. Consequently, genetic mutants that preferred targeting the most nutritious type of prey were directly favored by individual selection because they always had a resident who accepted to follow them since she was indirectly influenced by their mutated gene. The suboptimal coordination equilibrium was no longer an evolutionary trap. In finding a solution to the *behavioral* coordination problem, individuals solved the *evolutionary* coordination problem as well.

Put another way, behavioral coordination strategies changed the nature of the game. Individuals initially played a *coordination* game in which two players needed to jointly evolve a compatible preference. This raised a bootstrapping problem and made the transition from one equilibrium to another unlikely. By evolving endogenously a coordination strategy, individuals turned this game into a plain *optimization* game in which a single player was simply selected to choose the best possible prey.

Beyond the particular setting considered in this paper, we think these results reveal a general principle that could play a role in all games with multiple equilibriums, that is, in all coordination games, but also in repeated games such as the repeated prisoner's dilemma (see, for instance, Aumann and Shapley, 1994; Boyd, 2006). As a rule, there are many reasons why the behavior of one's partners will vary in all these games, making it necessary for one to adapt plastically to this variability (Mcnamara et al., 2010). In our simulations, for instance, individuals evolved a coordination strategy to adapt to the precise location where their partner was heading, but the same principle should hold in other settings as well. Even though behavioral plasticity originally evolves merely to deal with partners' phenotypic variability, it also happens to generate an adaptive response in front of mutants. These mutants probably did not exist when behavioral plasticity evolved, but they nevertheless happen to trigger the exact same response. And because this response was originally meant to maximize efficiency, it is likely to do so with mutants, too, as our experiments illustrate. Hence, there is a

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

general reason why the plastic response of individuals to each other should often "change the rules of the game" and smoothen the transition to efficient collective behaviors.

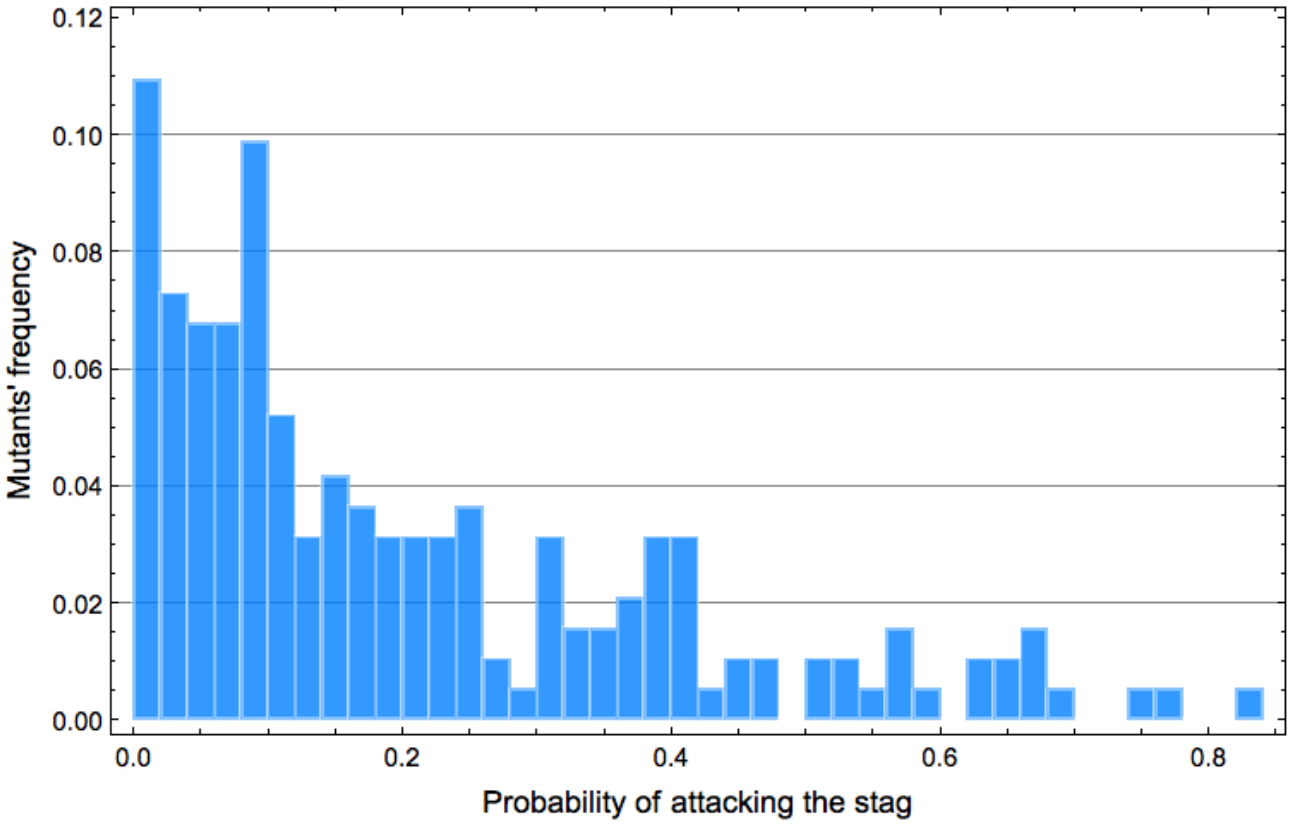

Figure 8: **Probability to prefer stags over boars in mutants generated randomly from a pure boar-hunter genotype.** Among the 10^5 mutants generated, we display only the 192 mutants whose probability to hunt stags is greater than 1% (see *Methods*).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4 Methods

4.1 Experimental setup

	Boar	Stag
Boar	125	0
Stag	0	250

Table 1: Possible outcome of a two-person coordination game. The game features two hunters. Each player may choose to hunt either a *boar*, or a *stag*. This payoff matrix has two equilibriums, a suboptimal one (both players hunt the boar) and an optimal one (both players hunt the stag).

The environment is a 800 by 800 unit arena with four solid walls. Each simulation is conducted with a collection of circular agents (with a diameter of 20 units). The agents are divided into two categories: hunters and prey. A pair of hunters and a varying number of prey (either 2 or 18) are always in the arena. The performance of the pair depends on the number of prey hunted as well as the type of those prey. There are two types of prey: boars and stags. The reward for boars is 125, while for stags it is 250 (a prey hunted alone otherwise yields no reward, Table 1). The initial positions of the prey are randomly determined at the beginning of each simulation, and the prey cannot move by themselves. To capture a prey, a robot has to move toward the prey and stay in contact with it for 800 time steps (out of a total of 20000 time steps for each simulation). Both robots have to be in contact with the prey at the end of the 800 time steps for the hunt to be considered successful. Once captured, the prey is removed from its location and replaced at a random position in the arena. We thus ensure that both the number of prey and the ratio of types of prey are kept constant.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hunters are equipped with two independent wheels and can move freely in the environment. They are also equipped with a set of sensors that are comprised of 12 proximity sensors and a camera. Proximity sensors are evenly distributed around the robotic agent's body, and each has a range of 40 units (that is, twice the hunter body's diameter). A proximity sensor corresponds to a ray toward a particular direction that indicates to the hunter the distance to the first obstacle (wall or other agent) in this direction. The camera is placed on the front of the robotic agent, and its 90 degree field of view is divided into 12 equally spaced rays. Each ray indicates the type (that is, hunter, boar, or stag) and the proximity of the nearest agent in its direction. Each hunter begins the simulation next to its partner at the top of the arena.

Robots are individually controlled by a fully connected multilayer Perceptron with a single hidden layer. The inputs of this neural network are fed with the sensory data of the hunters. One input neuron is used for each of the 12 proximity sensors, with maximal (resp. minimal) neural activity when the agent is directly in contact with an obstacle (resp. when there is no obstacle in the range of the sensor). Three neurons are used for each of the 12 rays of the camera: two neurons to encode the type of the agent in a two-bit binary value and one neuron to encode the proximity of the agent. Finally, there is a bias neuron whose value is always equal to one. Consequently, the number of input neurons totals 49. The hidden layer contains 8 neurons, while the output layer contains 2 neurons. These 2 output neurons control the speed of the left and right wheels; minimal (resp. maximal) activity results in maximal backward (resp. forward) actuation. Synaptic connections are each encoded in a single gene. Therefore, the total genome size is 410. Finally, the mapping function used to compute outputs is a sigmoid function.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4.2 Simulating artificial evolution

In each of the 30 independent replicates, we evolve a population of 20 individuals. Each individual is encoded as a genome, where each gene is a particular synaptic connection weight of the multilayer Perceptron controller. At the start of evolution, every gene in the genome is initialized with a random value sampled uniformly in $[0, 1]$. At each generation, the performance of every individual is evaluated by pairing this individual with five different partners chosen randomly in the population. The performance of each pair is evaluated during five trials to decrease the stochasticity effects that occur due to the random positioning of the prey. Hence, 25 simulations are performed for each individual per generation. The fitness score of an individual is computed as the average performance obtained throughout these 25 simulations, according to hunting performances (Table 1). After every individual is evaluated, we apply selection on the population using a $(10+10)$ elitist selection strategy Goldberg (1989). The population of the next generation is constituted from these 10 best (performance-wise) individuals of the current population and 10 new individuals, with each new individual being a mutated clone of a single parent (that is, there is no recombination, as is always the case with evolution strategies Beyer and Schwefel (2002)). Mutations are sampled according to a Gaussian operator, with a standard deviation of 2×10^{-1} and a per-gene mutation probability of 5×10^{-3}.

4.3 Complexity of the environment

To act on the complexity of the environment, we vary the number of prey present in the arena. Regardless of the number of prey, we ensure that the number of boars in the environment is always equal to the number of stags at any time of the simulation. Similarly, we also ensure that the number of prey stays constant throughout a simulation. In the "simple" environment, there are two prey in the arena: one boar and one stag. By contrast, in the "complex" environment, there are 18 prey: 9 boars and 9 stags.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4.4 Analyses on the transition between equilibriums

To determine whether the switch to the optimum collective equilibrium occurs at a given time, we compute the mean proportion of each type of prey hunted at each generation of evolution. We consider that if the mean proportion of stags hunted over the total number of prey hunted is greater than 0.5, then individuals in the population have switched to stag hunting.

4.5 Duplication and coevolution of neural networks

To study the impact of asymmetry between the hunters on their hunting strategy, we add the possibility for the individuals to coevolve two neural networks. While at the start of evolution every genome encodes for a single neural network, random duplication and deletion events could occur on offsprings' genomes after they have been mutated. When a duplication event takes place, each gene in the genome (with an original size of 410 genes) is duplicated to create a new genome (whose new size is thus 820). As each gene encodes for a synaptic weight in the neural network, this new genome thus encodes for two separate neural networks. When a deletion event occurs, a random half of the genome is deleted so that the genes encoding for one of the two neural networks are removed. The genome size thus reverts back to 410 genes. Each offspring's genome has a probability of 5×10^{-2} to be duplicated and a probability of 5×10^{-3} to be deleted at each generation. Duplications (resp. deletions) can only happen on nonduplicated (resp. duplicated) genomes.

Even after a duplication event, a given hunter is always controlled by a single neural network during the simulations. To study the effect of asymmetry, if only one of the two hunters possessed two neural networks, then this hunter would randomly choose one of those two neural networks.

We also conduct a control experiment to check that the evolutionary transition to stag hunting is facilitated by the leader-follower strategy rather than by the duplication of the neural network *per se*. In this control

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

experiment, duplication can still occur, but interacting individuals are always forced to use the same, randomly chosen, neural network, which prevents a division of labor from evolving. Under this treatment (i) individuals evolve a turning strategy rather than a leader-follower strategy, and (ii) the transition to stag hunting only occurs in 16 replicates out of 30 (Figure 7). This confirms that the division of labor, and not network duplication per se, facilitates the transition to stag hunting.

4.6 Analyses of boar-hunter mutants w.r.t. stag hunting

We generate 10^5 random mutants from a well-optimized boar-hunter genotype (with the same mutation parameters as in our evolutionary simulations), and assess each mutant's hunting preferences. From among these 10^5 mutants, we extract 192 mutants that displayed a probability greater than 0.01 of hunting the stag. Figure 8 shows the distribution of the preferences of these 192 mutants. Most mutants have only a small probability to hunt stags. In particular, not a single pure stag hunter can be found among the 10^5 mutants.

Acknowledgments: This work is supported by the European Union's Horizon 2020 research and innovation program under grant agreement No 640891 (DREAM project). Experiments presented in this paper were carried out using the Grid'5000 experimental testbed, being developed under the INRIA ALADDIN development action with support from CNRS, RENATER and several universities as well as other funding bodies (see https://www.grid5000.fr).

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Alvard, M. S., 2003. Kinship, Lineage, and an Evolutionary Perspective on Cooperative Hunting Groups in Indonesia. Human Nature 14:129–163.

Alvard, M. S. and D. A. Nolin, 2002. Rousseau ' s Whale Hunt ? 43:533–559.

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

Bernard, A., J.-B. AndrÃl', and N. Bredeche, 2016. To cooperate or not to cooperate: Why behavioural mechanisms matter. PLoS Comput Biol 12:1–14. URL `http://dx.doi.org/10.1371%2Fjournal.pcbi.1004886`.

Beyer, H.-G. and H.-P. Schwefel, 2002. Evolution strategies – a comprehensive introduction. Natural Computing 1:3–52. URL `https://doi.org/10.1023/A:1015059928466`.

Boyd, R., 2006. Reciprocity: You have to think different.

Boyd, R. and P. J. Richerson, 1990. Group Selection among Alternative Evolutionarily Stable Strategies. Journal of Theoretical Biology 145:331–342.

———, 2002. Group beneficial norms can spread rapidly in a structured population. J Theor Biol 215:287–296.

———, 2009. Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior. Journal of Theoretical Biology 257:331–339.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems 2:303–314. URL `https://doi.org/10.1007/BF02551274`.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Doncieux, S., N. Bredeche, J.-B. Mouret, and A. Eiben, 2015. Evolutionary robotics: what, why, and where to. Frontiers in Robotics and AI 2.

Drea, C. M. and A. N. Carter, 2009. Cooperative problem solving in a social carnivore. Animal Behaviour 78:967–977. URL http://linkinghub.elsevier.com/retrieve/pii/S000334720900339X.

Foster, D. and P. H. Young, 1990. Stochastic Evolutionary Game Dynamics. Theoretical Population Biology 38:219–232.

Goldberg, D. E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning.

Hughes, A. L., 2002. Adaptive evolution after gene duplication. Trends in Genetics 18:433–434.

Kandori, M., G. Mailath, and R. Rob, 1993. Learning, Mutation, and Long Run Equilibria in Games. Econometrica 61:29–56. URL http://www.jstor.org/stable/2951777.

Mcnamara, J. M., O. Leimar, and P. T. R. S. B, 2010. Variation and the response to variation as a basis for successful cooperation. Philosophical Transactions of the Royal Society B-Biological Sciences 365:2627–2633.

Mitri, S., S. Wischmann, D. Floreano, and L. Keller, 2013. Using robots to understand social behaviour. Biological reviews of the Cambridge Philosophical Society 88:31–9. URL http://www.ncbi.nlm.nih.gov/pubmed/22816672.

Nolfi, S. and D. Floreano, 2000. Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines. MIT Press. URL http://gral.istc.cnr.it/nolfi/er-book.html.

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

Richerson, P. J. and R. Boyd, 2005. Not by genes alone : how culture transformed human evolution. University of Chicago Press, Chicago.

bioRxiv preprint doi: https://doi.org/10.1101/422287; this version posted September 21, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams, 1986. Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chap. Learning Internal Representations by Error Propagation, Pp. 318–362. MIT Press, Cambridge, MA, USA. URL http://dl.acm.org/citation.cfm?id=104279.104293.

Trianni, V., 2014. Evolutionary Robotics: Model or Design? Frontiers in Robotics and AI 1:1–6. URL http://journal.frontiersin.org/article/10.3389/frobt.2014.00013/abstract.

Wolf, J. B., E. D. B. III, J. M. Cheverud, A. J. Moore, and M. J. Wade, 1998. Evolutionary consequences of indirect genetic effects. Trends in Ecology and Evolution 13:64 – 69. URL http://www.sciencedirect.com/science/article/pii/S0169534797012330.

Young, H. P., 1993. The Evolution of Conventions. Econometrica 61:57. URL http://www.jstor.org/stable/2951778?origin=crossref.