



HAL
open science

The mechanics of coordination and the evolution of cooperation : from computational modeling to evolutionary robotics design

Arthur Bernard

► **To cite this version:**

Arthur Bernard. The mechanics of coordination and the evolution of cooperation : from computational modeling to evolutionary robotics design. Robotics [cs.RO]. Université Pierre et Marie Curie - Paris VI, 2016. English. NNT : 2016PA066696 . tel-01609841

HAL Id: tel-01609841

<https://theses.hal.science/tel-01609841>

Submitted on 4 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A PhD Thesis by

Arthur BERNARD

submitted to

**Sorbonne Universités, Université Pierre et Marie
Curie**

École doctorale Informatique, Télécommunications et Électronique (Paris)

Institut des Systèmes Intelligents et de Robotique

in partial fulfillment of
the requirements for the degree of

Doctor in Computer Science

The Mechanics of Coordination and the Evolution of Cooperation

From Computational Modeling to Evolutionary Robotics Design

28/11/2016

Committee

Guillaume Beslon – LIRIS - Univ. Lyon/CNRS, FR	Reviewer
Richard Watson – Institute for Life Sciences - Univ. of Southampton, UK	Reviewer
Aurélie Beynier – LIP6 - UPMC/CNRS, FR	Examiner
Pierre-Henri Gouyon – Institut de Systématique, Évolution, Biodiversité - Sorbonne Universités, FR	Examiner
Laurent Keller – Department of Ecology and Evolution - Univ. of Lausanne, CH	Examiner
Jean-Baptiste André – ISEM - Univ. Montpellier, FR	Co-Advisor
Nicolas Bredeche – ISIR - UPMC/CNRS, FR	Co-Advisor

Address

Institut des Systèmes Intelligents et de Robotique
Université Pierre et Marie Curie
ISIR - CNRS UMR 7222
Boite courrier 173
4 Place Jussieu
75252 Paris cedex 05 - France

Keywords

Evolution of Cooperation, Coordination, Evolutionary Robotics, Stag Hunt, Multirobot Systems

Abstract

Many different species behave in a cooperative fashion and cooperation is central to most of the major transitions in evolution. However, explaining its evolution is a major challenge in evolutionary biology. The evolution of altruistic actions in particular, where individuals pay a cost to bring benefits to others, has been widely studied. In comparison, mutually beneficial actions, which benefit every individual participating, have been relatively ignored. Yet, while this type of cooperation is stable once evolved, explaining its origin is a challenge, in particular when it requires the coordination of several individuals.

In order to study the evolution of mutualistic cooperation, it is classical to use the evolutionary game theoretical model of the stag hunt. In this game, cooperating is more rewarding than acting in a solitary fashion but it is risky when rare. The issue is thus to study the emergence of the cooperative equilibrium. Here we claim that classical models in evolutionary biology make critical assumptions about the mechanics of behaviour that may impact the emergence of mutualistic cooperation. In consequence, we choose to address this issue with a framework that allows to take these mechanics into account: evolutionary robotics.

The fields to which we contribute in this thesis are twofold. First we use evolutionary robotics to model the evolution of mutualistic cooperation. Taking inspiration from the game of the stag hunt, we design an experiment of collective hunting. We show that while the transition to cooperation is easy in a classical game theoretical model, this transition becomes impossible with our model in evolutionary robotics. We thus reveal how modeling the practical mechanics of behaviours impacts the emergence of mutualistic actions. Then we show how individual selection alone may optimize collective actions as the emergence of coordination allows the transition to the optimum. Additionally, we reveal that the nature of the coordination behaviours evolved impacts the probability for this transition to occur.

In a second Part, we focus on the automatic design of controllers for distributed multi-robot systems. More precisely, we study the influence of genetic team composition on the design of cooperative agents in evolutionary robotics. We first compare a clonal approach (i.e. homogeneous team) and two aclonal approaches (i.e. heterogeneous team) in a collective foraging task. We reveal the existence of a tradeoff between the capacity to evolve cooperation, best achieved with homogeneous robots, and the efficiency of the cooperative solutions, where the more efficient cooperators are evolved with a particular aclonal approach: cooperative coevolution. Then we focus on the issue of evolving specialisation among heterogeneous robots. We study how specialisation can evolve at the level of the population, i.e. genotypic polymorphism. We reveal the critical challenges raised by this issue and that for genotypic polymorphism to occur, it is necessary to protect against the invasion of generalists as well as maintain sufficient genetic diversity in the population.

In conclusion, we show in this thesis how evolutionary robotics can contribute to a same problem (in our case the evolution of cooperation) in two very different directions: towards modeling in evolutionary biology or the automatic design of robots.

Contents

List of Figures	13
List of Tables	17
1 General Introduction	19
1.1 The Evolution of Cooperation	19
1.2 Model and Method	27
1.3 Evolving Coordination in Evolutionary Robotics	36
I Modeling the Mechanics of Coordination in the Evolution of Cooperation	39
2 Models in Evolutionary Biology and Evolutionary Robotics	41
2.1 Classical Models in Evolutionary Biology	42
2.2 Individual-Based Modeling	46
2.3 Evolutionary Robotics	48
2.4 Conclusions	51
3 The Impact of Behavioural Mechanisms in the Evolution of Cooperation	53
Abstract	55
Introduction	55
Materials and Methods	57
Results	60
Discussion	64
Supporting Information	68
4 The optimization of Collective Actions by Individual Selection	71
Abstract	73
Introduction	73
Materials and Methods	75
Results	76
Discussion	82

II	Designing Cooperative Robots in Evolutionary Robotics	87
5	Multirobot Systems and Automatic Design in Evolutionary Robotics	89
5.1	Multirobot Systems	90
5.2	Designing the Control of Collective Robots	94
5.3	Evolutionary Design for Distributed Robotics	98
5.4	Genetic Team Composition and the Evolution of Cooperation	101
5.5	Conclusions	105
6	The Tradeoff between Evolvability and Efficiency in the Evolution of Cooperation	107
	Abstract	109
	Introduction	109
	Methods	110
	Cooperative Foraging Task	112
	Going Beyond the Evolvability vs. Efficiency Tradeoff using Incremental Evolution	115
	Discussion and Conclusion	120
7	The Evolution of Specialisation through Genotypic Polymorphism	123
	Abstract	125
	Introduction	125
	Methods	126
	Behaviours of Specialists in a Cooperative Foraging Task	128
	Evolving Heterogeneous Behaviours with an Elitist Selection	130
	Evolution Under a Fitness-Proportionate Selection	132
	Computational Analyses of Population Dynamics	134
	Key Properties for Evolving Heterogeneous Behaviours	136
	Discussion and Conclusions	137
8	Discussion	139
8.1	Modeling the Evolution of Cooperation	139
8.2	Automatic Design of Collective Robots	143
8.3	Concluding Remarks	146
	Publications	149
	Bibliography	151



List of Figures

1.1 **Multiple examples of cooperation.** (A) Eusocial insects like Hymenoptera (e.g. ants) are capable of highly developed cooperative behaviours. (B) Some social carnivores like the spotted hyena perform various cooperative behaviours and are capable of advanced coordination strategies during collective hunts. (C) Social grooming is a cooperative behaviour where individuals groom each other reciprocally (here shown in impalas). (D) Interspecific cooperation (or mutualism) is a common cooperative behaviour. Here we show a *Labroides dimidiatus* involved in a cleaning symbiosis where the individual cleans another from parasites. 21

1.2 **Payoff matrix of the stag hunt.** In the stag hunt (Skyrms 2004), we consider that while hunting, two hunters can either hunt a hare or a stag. Hunting a hare can be done in a solitary or cooperative fashion, which ensures that any individual which hunts gets a reward. In comparison, hunting a stag can only be achieved in a cooperative fashion but rewards more than a hare. In consequence, an individual who would hunt a stag alone would not get any benefit. Payoffs are indicated in pair as follows: (Payoff for hunter 1; Payoff for hunter 2). The exact payoffs do not represent the most important aspect of the game as long as the different situations are in that order: R (reward for cooperation) > T (temptation for defection) = P (punishment for defection) > S (sucker's payoff). The payoff-dominant equilibrium is the equilibrium where the hunters maximize their maximum payoff whereas the risk-dominant equilibrium is the one where they maximize their minimum payoff. 28

1.3 **General workflow of an evolutionary robotics algorithm.** The main goal of ER is to evolve a population of genotypes. To that end, each genotype must be evaluated to obtain a fitness score. A genotype is thus translated into a phenotype (here an artificial neural network) and then embedded into a robot to act as its controller. The robot is situated in its environment and its behaviour is evaluated in accordance to the specificities of the task. Once every genotype has been assigned a fitness score, they undergo an evolutionary algorithm. This process selects the genotypes deemed fit to create offspring, on which variation is then applied. Finally this new population of genotypes replace the previous population and the process can go on for a new generation. 30



1.4	Robot model of our experimental setting. This figure represents the sensory and neural architecture of the simulated robotic agents used in our experimental study. On the robot diagram, proximity sensors are represented by the blue lines whereas the front camera is shown as a red cone. The neural network is a multilayer perceptron with one hidden layer and whose inputs are constituted of all the sensory information of the individual. The outputs of the neural networks are the speeds of both of the robot's wheels.	34
3.1	Screenshot of a robotic simulation. The red dots represent the two hunters, the green dots the hares, and the pink dots the stags. The black lines around the agents' body represent the proximity sensors and the black cones on front the cameras described in the text. Hunters are allowed to move throughout the environment. Hares and stags remain at their starting positions.	57
3.2	Diagram of the simulated robotic agent used in the simulation (inset) and its neural network controller. The blue lines represent the 12 proximity sensors and the red lines represent the front camera. Inputs "Type 1" and "Type 2" are two boolean values used to represent the type of the agent (encoded with two bits) recognized by the camera ray.	59
3.3	Evolution of cooperation in a robotic simulation with an initial hare-hunting strategy. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted successfully (cooperatively), and the top grey bar the number of failed hunts (stags hunted alone). The standard deviation for each quantity is shown by black lines. The population for each of the 30 independent runs was previously evolved in an environment with only hares. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.	61
3.4	Evolution of cooperation in a game-theoretic simulation with an initial hare-hunting strategy. Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted when starting with a population of hare hunters for 30 independent runs. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1.	62

3.5	<p>Evolution of cooperation with no initial hunting strategy. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted successfully (cooperatively) and the top grey bar the number of failed hunts (stags hunted alone). The standard deviation for each quantity is shown by black lines. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively, as presented in Table 3.1. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.</p>	63
3.6	<p>Evolution of cooperation with an initial hare-hunting strategy and a reward for solitary stag hunting. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted cooperatively and the top grey bar the number of stags hunted alone. The standard deviation for each quantity is shown by black lines. The population for each of the 30 independent runs was previously evolved in an environment with only hares. Rewards were 50 for a hare, 50 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.3. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.</p>	65
3.7	<p>Evolution of cooperation under maximal genetic relatedness with an initial hare-hunting strategy. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted successfully (cooperatively) and the top grey bar the number of failed hunts (stags hunted alone). The standard deviation for each quantity is shown by black lines. The population for each of the 30 independent runs was previously evolved in an environment with only hares. The genetic relatedness between paired individuals was 1. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.</p>	66

3.8	Snapshots of a simulation after two hunts. In each of these snapshots, we show the path travelled by each hunter (in different colours) since their last prey was hunted. The black dots represent the positions of the hunters at their last kill. The red star on the stag (pink circle) converged on by the hunters indicates not only that the prey was killed but, more importantly, that it was killed cooperatively by the two hunters.	67
3.9	Evolution of cooperation in a game-theoretic simulation with an initial hare-hunting strategy. Evolution of the mean percentage of stags hunted with respect to the total number of prey hunted where individuals are initially unable to hunt for 30 independent runs. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively.	68
3.10	Evolution of cooperation under maximal genetic relatedness with no initial hunting strategy. Evolution of the mean percentage of stags hunted with respect to the total number of prey hunted in a robotic simulation. The genetic relatedness between paired individuals was 1. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented.	69
3.11	Evolution of cooperation with an initial hare-hunting strategy and a varied density of prey. Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation when the number of prey was (A) 6 and (B) 30. The population for each of the 30 independent runs was previously evolved in an environment with only hares. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1. The number of prey was kept constant throughout the simulation by replacing killed prey by a prey of the same type.	69
4.1	Mean proportion of prey hunted cooperatively. (A) Mean proportion of boars and stags hunted cooperatively by the best individual in each of the 30 independent replications. The colored areas around the medians represent the first and third quartiles. The red line represents the separation between the pre-evolution step (when hunting stags rewards nothing) and the rest of the evolution. (B) Repartition of the prey hunted at the last generation of evolution by the best individual in each replication. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1).	77

4.2	Proportion of stag hunting runs and proportion of prey hunted. (A) Number of replications (out of a total of 30) where stag hunting evolved in the <i>Control</i> and <i>Coordination</i> settings. We consider that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the <i>Control</i> setting, the environment is constituted of one boar and one stag. In comparison, in the <i>Coordination</i> setting, 18 prey are present in the environment and it is thus necessary to coordinate for cooperation to happen. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1). (B) Repartition of prey hunted at the last generation of evolution by the best individual in every replication in the <i>Coordination</i> setting. The population for each replication previously evolved in an environment where hunting stags rewarded nothing.	78
4.3	Display of a turning strategy after an entire simulation. Both individuals adopt a <i>turning</i> strategy during a complete simulation. The paths of the agents are 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.	79
4.4	Proportion of cooperative runs. Number of replications (out of a total of 30) where stag hunting evolved in the <i>Control</i> , <i>Coordination</i> and <i>Coordination+Duplication</i> settings. We consider that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the <i>Control</i> setting, the environment is constituted of one boar and one stag. In comparison, in the <i>Coordination</i> setting, 18 prey are present in the environment and it is thus necessary to coordinate for cooperation to happen. Finally, in the <i>Coordination+Duplication</i> setting, the environment is also constituted of 18 prey but individuals have added neural plasticity which allows them to adopt more complex strategies and in particular a leader/follower strategy. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1). The population of each replication previously evolved in an environment where hunting stags rewards nothing.	81
4.5	Display of a leader/follower strategy after an entire simulation. Both individuals adopt a leader/follower strategy during a complete simulation. The paths of the agents are 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 blue 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.	82

4.6	Mean reward comparison of turning and leader/follower strategies. Mean reward of the best individuals in the 30 replications over evolutionary time by individuals adopting a turning strategy (evolved in the <i>Coordination</i> setting) or a leader/follower strategy (evolved in the <i>Coordination+Duplication</i> strategy) during the pre-evolution step. Rewards were 50 for a boar if hunted alone, 125 if hunted cooperatively, 0 for a stag hunted alone and 250 if hunted in a cooperative fashion (Table 4.1). . . .	83
4.7	Proportion of cooperative runs. Number of replications (out of a total of 30) where stag hunting evolved in the <i>Control</i> , <i>Coordination</i> and <i>Coordination+Duplication</i> settings. We consider that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the <i>Control</i> setting, the environment is constituted of one boar and one stag. In comparison, in the <i>Coordination</i> setting, 18 prey are present in the environment and it is thus necessary to coordinate for cooperation to happen. Finally, in the <i>Coordination+Duplication</i> setting, the environment is also constituted of 18 prey but individuals have added neural plasticity which allows them to adopt more complex strategies. At a beginning of a simulation, both individuals use the same neural network. A change of network occurs for an individual when the other individual gets on a prey. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1). The population of each replication previously evolved in an environment where hunting stags rewards nothing. . . .	85
4.8	Mean reward comparison of turning and search/join strategies. Mean reward of the best individuals in the 30 replications over evolutionary time by individuals adopting a turning strategy (evolved in the <i>Coordination</i> setting) or a search/join strategy (evolved in the <i>Coordination+Duplication</i> strategy) during the pre-evolution step. Rewards were 50 for a boar if hunted alone, 125 if hunted cooperatively, 0 for a stag hunted alone and 250 if hunted in a cooperative fashion (Table 4.1). . . .	86
6.1	Performance of the cooperative solutions. Median fitness score of the best individuals in each of the runs where cooperation evolved for each setup over time. The fitness score of an individual is computed as the average reward the individual earned per trial by foraging targets. The colored areas around the medians represent the first and third quartiles. . .	114
6.2	Snapshots of the simulation after an entire trial in the foraging task. The path of each robotic agent from their initial positions (black dots) is represented in red and blue. The green and purple discs represent the 18 targets in the environment. When a target is foraged by the two agents, a red cross (resp. blue) is drawn on the target if the red agent (resp. blue) arrived on it first. Each snapshot corresponds to a trial where agents adopted a different behavior: (A) turning or (B) leader/follower. . .	115

6.3	Performance of the cooperative solutions. Median fitness score of the best individuals in each of the 60 independent runs and for each setup over time. Fitness score is computed as the average longest sequence of waypoints shared by both agents per trial. The colored areas around the medians represent the first and third quartiles.	117
6.4	Proportion of leadership. Boxplots of the proportion of leadership over time for the best individuals in each runs where the proportion at the last evaluation was greater than 0.75 in the (A) control, (B) clonal or (C) coevolution setup. This value represents the proportion of waypoints crossed by both individuals for which the leader arrived first.	119
6.5	Performance of the cooperative solutions. Median fitness score of the best individuals in each of the runs where cooperation evolved for each setup over time. The fitness score of an individual is computed as the average reward the individual earned per trial by foraging targets. The colored areas around the medians represent the first and third quartiles.	120
7.1	Snapshots of the simulation after an entire trial in the foraging task. The path of each robotic agent from their initial positions (black dots) is represented in red and blue. The blue discs represent the 18 targets in the environment. When a target is foraged by the two agents, a red cross (resp. blue) is drawn on the target if the red agent (resp. blue) arrived on it first. Each snapshot corresponds to a trial where agents adopted a different strategy: (A) turning or (B) leader/follower.	128
7.2	Average reward and leadership proportion with a leader/follower or turning strategy Boxplots of (A) the average reward and (B) the leadership proportion over 20 independent trials for the leader/follower and turning strategies. The leadership ratio of an individual represents the propensity for one individual among the pair to arrive first more often than its partner on a target collected in a cooperative fashion. The position of each target at the beginning of each trial was randomized.	129
7.3	Proportion of time with a leader/follower strategy. Boxplots of the number of generations where the best individual in each replicate adopted a leader/follower strategy out of the total number of generations. We consider that the best individual adopted a leader/follower strategy when its leadership ratio was over a threshold value of 0.6.	131
7.4	Vector field of the gradient of phenotypes' proportions and proportions of phenotypes at last generation of evolution. (A) Vector field of the gradient of phenotypes' proportions in an infinite population. The strength of variation is indicated by the color of the arrow. (B) Repartition of phenotypes at the last generation of evolution for all three population sizes. Evolution lasted 1500 generations and results were replicated across 11 independent simulations. The initial population was entirely composed of <i>leaders</i>	135

List of Tables

1.1	Experimental parameters.	33
2.1	Payoff matrix of the prisoner’s dilemma. The strategy of player A (resp. player B) is symbolized by each row (resp. column). The payoff of player A (resp. player B) is shown on the left (resp. right). This is the payoff matrix which was used by Axelrod & Hamilton in their work on Tit for Tat in the prisoner’s dilemma (Axelrod 1984).	44
3.1	Food rewards for hunting different prey. The reward depends on whether these prey were hunted alone or cooperatively. There is no reward for stags hunted alone in this case.	58
3.2	Simulation parameters.	60
3.3	Food rewards for hunting different prey. The reward depends on whether these prey were hunted alone or cooperatively. There is a reward for stags hunted alone in this case.	64
4.1	Food Rewards for hunting. Rewards depend on whether the hunt was solitary or cooperative.	76
5.1	Architecture choices in multirobot systems.	92
6.1	Rewards for the foraging of the different targets. Rewards depend on whether they were collected alone or cooperatively.	112
6.2	Evolution of a cooperative strategy. Number of simulations where the best individual evolved a cooperative strategy (collecting purple targets) or a solitary strategy (collecting green targets) for each setup in the foraging task.	113
6.3	Evolution of a cooperative strategy. Repartition of the different strategies evolved in each of the runs where cooperation evolved for each setup in the foraging task. We indicate in each cell the number of simulations where a particular strategy evolved.	116

6.4 **Cooperative strategies evolved.** Repartition of the different strategies evolved in each of the 60 independent runs for each setup in the way-points task. We indicate in each cell the number of simulations where a particular strategy evolved: *Leader/follower* (Lead.), *Turning* (Turn.) or *Other*. “Other” regroups wall-following strategies or simulations where no recognizable strategy evolved. 118

6.5 **Evolution of a cooperative strategy.** Proportion of the 60 independent simulations where the best individual evolved a cooperative strategy (collecting purple targets) or a solitary strategy (collecting green targets) for each setup in the foraging task when individuals are previously evolved in the waypoints task. In addition, the repartition of the different strategies is indicated when cooperation evolved: *Leader/Follower* (Lead.) or *Turning* (Turn.). 118

7.1 **Rewards for the foraging of targets.** Rewards depend on whether the targerers were collected in a solitary or cooperative fashion. 127

7.2 **Strategies evolved by the best individuals under elitist selection with an initially random population.** Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the replicates for different population sizes N . We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under an $(\mu + \lambda)$ elitist selection, with $\mu = \frac{N}{2}$ and $\lambda = \frac{N}{2}$. Individuals’ genotype values were intially random. In the table “L/F” stands for leader/follower and “NC” for “Non-Cooperative”. 130

7.3 **Strategies evolved by the best individuals under elitist selection when adding followers.** Repartition of the different strategies adopted by the best individuals at last evaluation in each of the replicates for different population sizes N . We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under a $(\mu + \lambda)$ elitist selection, with $\mu = \frac{N}{2}$ and $\lambda = \frac{N}{2}$. The population was initially seeded with a population of leaders in which we added a specific amount of followers. In the table “L/F” stands for leader/follower and “NC” for “Non-Cooperative”. 132

7.4 **Strategies evolved by the best individuals under fitness-proportionate selection with an initially random population.** Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the replicates for different population sizes. We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under a fitness-proportionate selection. Individuals’ genotype values were initially random. In the table “L/F” stands for leader/follower and “NC” for “Non-Cooperative”. 133

7.5	Strategies evolved by the best individuals under fitness-proportionate selection when adding followers. Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the replicates for different population sizes N . We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under a fitness-proportionate selection. The population was initially seeded with a population of leaders in which we added a specific amount of followers. In the table "L/F" stands for leader/follower and "NC" for "Non-Cooperative".	133
7.6	Payoff matrix for pair-wise simulations of each phenotype. Average payoffs of each phenotype against every phenotype in a pair-wise simulation. Each pair was evaluated 10 times in order to decrease the stochastic effects of the initial conditions (i.e. random positions of the targets).	134
7.7	Strategies evolved by the best individuals when coevolving two populations. Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the 11 replicates. We indicate in each cell the number of simulations where a particular strategy evolved. Two populations were coevolved under elitist selection and the individuals' genotype values were initially random. In the table "L/F" stands for leader/follower and "NC" for "Non-Cooperative".	136



Chapter 1

General Introduction

Natural selection cannot possibly produce any modification in any one species exclusively for the good of another species.

— Charles Robert Darwin

1.1	The Evolution of Cooperation	19
1.1.1	The Problem of Evolving Cooperation	19
1.1.2	Altruism and Indirect Fitness Benefits	22
1.1.3	Direct Fitness Benefits and Mutualism	23
1.1.4	Stability Versus Origin	25
1.1.5	Proximate and Ultimate Mechanisms	26
1.2	Model and Method	27
1.2.1	Game Theory and the Stag Hunt	27
1.2.2	Evolutionary Robotics	29
1.2.3	Experimental Setting	32
1.3	Evolving Coordination in Evolutionary Robotics	36
1.3.1	Modeling the Evolution of Coordination	36
1.3.2	Designing Cooperative Robots	36
1.3.3	Organization of the Manuscript	37

1.1 The Evolution of Cooperation

1.1.1 The Problem of Evolving Cooperation

Many different species exhibit cooperative behaviours in very diverse manners. It is even argued that cooperation is one of the leading factors in most of the major transitions in evolution (e.g. the appearance of eukaryotes or the evolution of multicellularity) (Szathmàry and Maynard Smith 1995). The classical definition of cooperation in evolutionary biology is as follows: cooperation is a behaviour where an *actor* (the individual who initiates the behaviour) behaves in such a way that is beneficial to a *recipient* (West et

al. 2007b). Given how broad this definition is, numerous biological phenomena can be included under this label.

Cooperative actions are present at every level of the natural world. Even unicellular organisms such as bacteria and microorganisms are known to frequently act in a cooperative manner. By using secretions, microorganisms are capable of collective sharing and communication (Elena and Lenski 2003; Keller and Surette 2006; West et al. 2006). *Pseudomonas aeruginosas* for example produce nutrients that every organism in the vicinity can benefit from (Popat et al. 2012; Harrison 2013). Insects from the Hymenoptera (e.g. ants, wasps, bees) and Isoptera (e.g. termites) orders are known for the presence of *eusociality* (Figure 1.1 (A)) (Wilson 1990) which entails highly cooperative behaviours between individuals. In particular, the most distinctive feature of eusociality is the existence of division of reproductive labor. This means that reproductive and non-reproductive castes (e.g. worker caste) coexist, where individuals which cannot reproduce care for the young of others. Some social carnivores are capable of collective hunting, where several members of the same group coordinate their actions to catch a challenging prey. Spotted hyenas (see Figure 1.1 (B)) in particular are efficient hunters which rely on signaling and communication to coordinate (Drea and Carter 2009; Smith et al. 2010, 2012) and are able to defend their catch against lions. But they are also considered to be the most social taxon among Carnivora (Mills 2003) and the complexity of their social organization is comparable to that of primates (Drea and Frank 2003). Finally, the scope of cooperative behaviours is such that even cooperation between individuals from different species (i.e. interspecific cooperation or mutualism) is abundant (Bshary 2004). Cleaning symbiosis, where a "client" has its teeth or body cleaned from parasites or dead tissues by a smaller "cleaner" (Poulin and Grutter 1996), is an example of mutualism. In particular some fishes, especially *Labroides dimidiatus* (wrasses, Figure 1.1 (D)), are known to clean other bigger fishes to the point that there exists "cleaning station" where multiple aquatic animals converge to benefit from their services.

Yet explaining the evolution of cooperation has been one of the major challenges in evolutionary biology (Hamilton 1964; Dugatkin 2002; West et al. 2011). Charles Darwin had already said that the evolution of cooperation could pose a problem to his theory. He thought that the existence of a non-reproductive caste in eusocial insects was "one special difficulty, which at first appeared to me to be insuperable, and actually fatal to my whole theory" (Darwin 1859). According to the theory of evolution, life is a struggle where only the *fittest individuals* survive. The main purpose of evolutionary biology is to explain adaptation (West et al. 2011). In particular, natural selection is driven by the reproduction of individuals. Namely, because transmission of genetic material occurs through reproduction, evolution leads to an increase in genes and traits that increase the relative number of offspring (i.e. fitness) of the organism. This is where this understanding of adaptation appears to contradict with the evolution of cooperation.

Cooperation is defined as a behaviour that benefits another individual than the actor. One must keep in mind that "costs" and "benefits" refer to the fitness of the individual (i.e. the number of this individual's offspring) and not direct material elements. In consequence, a cooperative behaviour decreases the relative fitness of the actor compared to that of others in the population and should then be selected against. In particular,

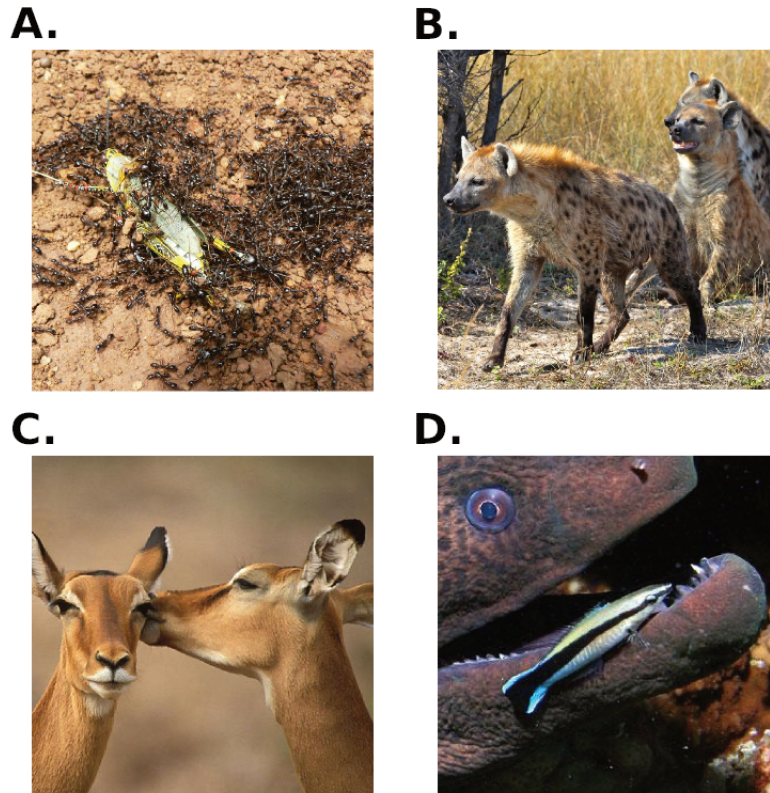


Figure 1.1 – Multiple examples of cooperation. (A) *Eusocial insects like Hymenoptera (e.g. ants) are capable of highly developed cooperative behaviours.* (B) *Some social carnivores like the spotted hyena perform various cooperative behaviours and are capable of advanced coordination strategies during collective hunts.* (C) *Social grooming is a cooperative behaviour where individuals groom each other reciprocally (here shown in impalas).* (D) *Interspecific cooperation (or mutualism) is a common cooperative behaviour. Here we show a *Labroides dimidiatus* involved in a cleaning symbiosis where the individual cleans another from parasites.*

cooperative individuals are under the threat of invasion by cheaters (or freeloaders¹). To illustrate this issue, let us take back the example of *P. aeruginosas* which we previously talked about. Assume that a population is constituted of cooperators which can pay a certain cost to give a benefit to others. A given individual can benefit from every individual in the population and the situation appears ideal. However, say now that a mutant which does not cooperate, appears in the population. Then this mutant can benefit from the cooperative actions of others without having to pay any cost. From this it stems that its relative fitness will be higher than that of cooperators. In consequence, it will produce more offspring and mutants will begin to invade the population. The process will repeat itself as cheaters have a higher fitness than cooperators until no cooperators

¹The terms "cheaters" and "freeloaders" will appear several times along the lines of this manuscript. Because this word has a bad connotation with regards to human actions, it may be misunderstood that we attribute a malevolent intention to the organisms. Rather, through the rest of the manuscript, it should be understood as "one that profits from a collective benefit without paying any cost".

are left in the population. From this comes that cooperation should not be able to exist.

Biologists have therefore been studying the mechanisms which could explain the evolution of cooperation for several decades. Because cooperative actions are so diverse, numerous models have been proposed to classify the different mechanisms which could explain the adaptation of cooperative traits (Dugatkin 2002; Keller and Surette 2006; Bergmüller et al. 2007; West et al. 2007a). For our overview of cooperation in this present manuscript, we follow the classification of West and colleagues (West et al. 2007b). In particular, these mechanisms can be classified in two main categories: *direct fitness benefits* and *indirect fitness benefits*.

1.1.2 Altruism and Indirect Fitness Benefits

A particular type of cooperative actions that has garnered a lot of attention is *altruism*. We consider a behaviour to be altruistic when the actor of the behaviour pays a fitness cost in an action benefitting another individual (Hamilton 1964; West et al. 2007b). The costly secretion of nutrients by *P. aeruginosas* corresponds, in its simplest form, to an altruistic action. Eusociality is also a major example of altruism in the natural world. In particular, the distribution of reproductive labour means that part of the individuals do not reproduce at all, thus paying the highest fitness cost possible.

The main problem posed by the evolution of altruism is its stability against the invasion of cheaters. As previously explained with the example of *P. aeruginosas*, cheaters can easily invade the population and take over cooperators. This sparked numerous research on the evolution of altruism. The now well-known explanation for the evolution of altruism was proposed by Hamilton: *kin selection* (Hamilton 1964). This mechanism conveys the idea that a particular trait can spread through the reproduction of relatives. If we consider the unit of selection in evolution to be the gene (as explained by Dawkins with the *selfish gene* metaphor (Dawkins 1976)), then the ultimate "goal" of a gene is to spread in the population. It thus does not really matter if a particular individual reproduces or not. If an individual helps another individual that is genetically close to her (i.e. genetically related), she is still transmitting similar genetic material even if she ultimately does not produce offspring. In consequences, an altruistic trait can still spread in the population through helping a relative who may bear this trait: this is kin selection. This general idea of the transmission of one's genes through a relative is encapsulated into the concept of *indirect fitness benefits*. Namely, if a trait contributes to the fitness of relatives, then this trait will also be favoured by natural selection. This wider definition of fitness is called the *inclusive fitness* (Grafen 1984), which is constituted of both the direct and indirect fitness benefits. As an example, it is now known that kin selection is a driving evolutionary mechanism for the evolution of eusociality. In particular empirical works on eusocial insects have shown that colonies are composed of strongly related organisms. Namely all members of the colony are offspring of a single individual (i.e. the queen) (Queller and Strassmann 1998; Bourke 2014).

While discussing altruism we want to address the long time debate about the influence of *group selection* on the evolution of altruistic behaviours. In particular, the evolution of cooperation was not deemed worthy of studying for some time because it was thought

that it could easily be explained by group selection (Axelrod and Hamilton 1981). There has been two major branches of this theory, now called old and new group selection theory (West et al. 2007b). In the old group selection, it was considered that a cooperative trait could evolve because it was beneficial to the whole group. For example, assume that there are two groups in competition, one constituted of cooperators and the other of defectors, who consume resources selfishly. Because defectors exploit their resources selfishly, they would go extinct when resources disappear. Thus the group of cooperators, and therefore cooperation, could survive because selection acts at the level of the group. This is why we may often hear the idea that an individual behaves in a certain way "for the good of the species". While this particular theory of group selection was then dismissed as nonexistent (Maynard Smith 1976), a new group selection theory arised. In this new theory, the main idea is that individuals interact in small groups, which exist inside a given population (whereas old group selection considered the whole population to be the group). Because interactions take place between a small number of individuals inside a given group, then the emergence of cooperative traits can be favored. Since then, it has been shown that kin selection and new group selection are mathematically identical concepts (Hamilton 1975; Van Baalen and Rand 1998; Gardner et al. 2007) and that it is generally easier to use the kin selection framework (West et al. 2007b). Still, it is argued that group selection might occur in the maintenance of sex through the effects of lineage selection (Nunney 1989; Vienne et al. 2013).

1.1.3 Direct Fitness Benefits and Mutualism

But not all cooperative actions are altruistic. In fact cooperation can also be directly beneficial to the actor (Leimar and Hammerstein 2010). In this case, both the actor and the recipient benefit from the cooperative behaviour and we then say that the behaviour is mutually beneficial (West et al. 2007b). Before we talk more thoroughly about this subject we must clarify some confusions that sometimes arise in the literature (Bergmüller et al. 2007). Some have considered cooperation to only refer to mutually beneficial behaviours (Trivers 1985; Lehmann and Keller 2006) in comparison to altruism. Here we consider the broader definition where cooperation includes both altruistic and mutualistic actions (West et al. 2007b). There is also some confusion about the definition of *mutualism*. While it can be used to describe mutually beneficial actions or sometimes even cooperation as a whole (as previously explained), it may also strictly refer only to interspecific mutualism. In the context of this thesis, we are mainly interested in intraspecific cooperation. As such, we follow the advice of West et al. (West et al. 2007a) and use "mutually beneficial behaviours" rather than mutualism throughout this manuscript.

Different mechanisms have been proposed in order to explain how cooperation can be adaptive through direct benefits. For example the benefits for the actor can be enforced in multiple manners. An exhaustive review of enforcement is beyond the scope of this manuscript but we will quickly describe this mechanism to give a general overview of its influence. First, one way those benefits can be enforced is through reciprocal interactions (Trivers 1971). Under reciprocity, individuals will tend to help those who have helped them in the past and thus provide mutual (albeit delayed) benefits. In this case, we

talk about direct reciprocity. In comparison, under indirect reciprocity an individual will tend to help an individual who is known to help others, hence the notion of "reputation-based reciprocity". It is interesting to note that, outside humans, reciprocity is thought to be generally unimportant (Dugatkin 1997). However, one well-known example of non-human reciprocity is in the allogrooming behaviour of impalas (see Figure 1.1 (C)) (Hart and Hart 1992). Impalas groom each other in order to remove ticks from the other individual. What is particularly interesting is that grooming occurs in several alternative bouts where one individual bouts the other. More importantly, each individual is actor and recipient for the same number of bouts. This behaviour is widespread and can involve pairs of males, females and fawns. Furthermore, the individuals are unlikely to be related, which removes kin selection as a possible explanation.

Other forms of enforcement (or coercion (Clutton-Brock 2002)) include rewards, punishment, sanctions and policing. These types of enforcement are common in humans (Fehr and Gächter 2002) but also in a lot of different social animals. Spotted hyenas (Drea and Frank 2003; Drea and Carter 2009; Smith et al. 2012) enforce cooperation inside the group through suppressed reproduction. Dominant females may attack lower ranking females if they get pregnant or even attack their cubs directly. They thus ensure that their offspring are the only young in the group. This way, they enforce cooperative care (alloparenting) of young and increase the survival chances of their offspring. The cleaner fishes we previously talked about provide an elegant way to distinguish between various enforcement mechanisms. While they eat parasites from their client, their preference is to eat the mucus and tissue. In return, clients use different ways to solve this conflict and enforce cooperation (i.e. the cleaner must only eat parasites): partner choice, which means that they only go to "good" cleaners (i.e. cooperators), partner switching, which means that they choose to be cleaned by another individual, and directly punishing the cleaner through aggression (Bshary and Grutter 2005). Additionally, in the case of wrasses this means that there are specific mechanisms which sustain this mutualism. In consequence, this implies that there is a particular investment in this behaviour for the specific benefits of the mutualistic relation.

In the context of this thesis we study the case where cooperation between individuals is not enforced. Rather, all individuals have a shared interest in cooperating. This is something which is often called *by-product benefits*, conveying the idea that cooperation is a by-product from an otherwise self-interested action. For example, a large group of individuals entails higher chances of survival (against both the environment and predators) and an increase in the benefits from foraging or hunting. This leads individuals to have a mutual benefit in creating groups and societies, something which is known as group augmentation (Bergmüller et al. 2007) and has been well studied, for example in the case of meerkats (Clutton-Brock 2002). But by-products can also lead to high degrees of coordination between individuals through the benefits of coordinated by-products (Leimar and Connor 2003). In this case, individuals act in their own interest but also react to the behaviour of others. In particular, cooperative hunting is a very significant type of coordinated by-products. For instance, we previously talked about spotted hyenas which display complex and very coordinated hunting strategies.

An example of the selfishness involved in by-products was provided by Caraco & Brown

with the Mexican Jay (Caraco and Brown 1986; Dugatkin 2002). When there is food in large enough quantity, these birds will share food with other individuals' offspring. While this behaviour may appear altruistic at first, it can be explained by purely selfish motives. In particular, chicks beg loudly until they are fed, which might attract nearby predators. Thus, an individual may share food so that another individual's chicks do not attract the attention of predators to its own offspring. This example illustrates how it can be easy to confuse purely selfish behaviours for altruistic ones. Similar observations have been made from the behaviours of sentinels, where individuals will take the role of watching for predators in order to alert others in the group. One could think that sentinels act in an altruistic manner as they do not forage while they stand guard and might attract the predator's attention to themselves when signaling to others. Yet in the case of the Arabian babbler (Wright et al. 2001), an individual goes on sentinel duty only under certain conditions. In particular, individuals will act as watchers when they have already collected enough food to satisfy their needs. Being a sentinel is then a way to increase their own survival chances against predators. Similarly, meerkats selfishly benefit from standing guard (Clutton-Brock et al. 1999). Additionally, it was shown that there exists no evidence that they take a higher risk to be killed by a predator when acting as sentinels.

Finally, as a conclusion, it is important to state that the differences between all these mechanisms are generally subtle and challenging to differentiate. In particular, indirect and direct fitness benefits are not necessarily mutually exclusive and a behaviour can evolve thanks to kin selection but then its benefits may be generated through enforcement. Mutually beneficial actions are prevalent in among social actions and thus the question of their evolutionary origin is critical.

1.1.4 Stability Versus Origin

When interested in the adaptation of cooperation, we can distinguish between two different issues: *stability* and *origin*. In the case of altruistic behaviours, the stability of cooperation is under the constant threat of invasion by freeloaders. As such the main problem involved is to study their stability against subversion by cheaters. In comparison, when benefits are mutual there does not seem to be any evolutionary puzzle in the same way as with altruism: once a mutualistic behaviour has evolved, there is no incentive to free-ride (Forber and Smead 2015). Thus the main focus for mutualistic behaviours is not to understand the stability of cooperation but its origin (West et al. 2007a).

In particular, there is an issue raised by the evolution of mutually beneficial behaviours when they require that individuals coordinate (Alvard and Nolin 2002; Alvard 2003; Leimar and Connor 2003; Drea and Carter 2009). For example, the evolution of collective hunting is one that is based on interdependence between the individuals (Tomasello et al. 2012), which means that the collective success of the group is dependent on the coordination between several individuals. Because of this interdependence, the problem is not one of stability as there is no selective advantage in cheating. However there is a fitness valley to cross where all the individuals have to evolve coordination before they are able to reap the benefits of collective actions. To put it more simply, we face a *chicken and egg dilemma*. For a cooperative trait to be selected, it needs to benefit the individual.

In this case this benefit cannot be obtained unless others are also able to coordinate. In consequence, the other individuals must have already evolved cooperation. There has thus been a recent shift in evolutionary biology toward the study of the origin of cooperation (in the case of mutualistic actions) rather than its stability (Forber and Smead 2015).

1.1.5 Proximate and Ultimate Mechanisms

There are two ways in which we can approach the study of animal behaviours and this distinction is critical in the context of this thesis. These two approaches were introduced by Niko Tinbergen (Tinbergen 1963; West et al. 2007b) as complementary manners in the way we look for evolutionary explanations of behaviours :

- We can study the mechanics of behaviour to give *proximate* explanations
- We can be interested in the fitness consequences of the behaviour and reveal the *ultimate* explanations

To put it more simply, we can abstract from the practical interactions that take place and focus on explaining *why* a particular behaviour is adaptive, which is an ultimate explanation. Or we can consider *how* the behaviour functions and thus be interested in the proximate mechanisms. Tinbergen illustrated this difference with the example of the black-headed gull. These birds remove the eggshells from their nest. The proximate (or mechanistic) explanation is that individuals will more likely remove objects from the nest when they have frilly edges and are egg-coloured and feather-light. The ultimate explanation is that predators are this way less likely to spot their offspring. The conclusion is that both explanations are necessary so that we can fully understand this behaviour. Scott-Phillips and colleagues (Scott-Phillips et al. 2011) provided another way to summarize the difference between these two explanations: proximate mechanisms generate behaviours whereas ultimate functions explain why these behaviours are favored.

In this thesis, we want to show that when we are interested in the origin of cooperation rather than its stability it is necessary to study the proximate as well as the ultimate explanations. In particular, because we are interested in the origin of cooperative actions rather than their stability, proximate mechanisms are critical. These mechanisms indeed affect the availability of individual mutations. Namely, the possibility for particular mutants to appear is dependent on the nature of these mechanisms. When we are interested in the study of the stability of cooperation, the goal is to show that no mutant can invade and replace cooperators in the population. In consequence we do not focus on the manner in which these mutants may appear in the population because it is conservative to do so. However it is not conservative with regards to the emergence of cooperative actions. Studying the origin of cooperation implies that it is necessary to prove that there exists a gradual convergence toward a cooperative behaviour. Therefore assumptions about the appearance of mutants may be critical to the evolution of cooperation. In particular, we will show in Chapter 2 that one such assumption made by classical models in evolutionary biology is to consider that the effect of mutations are small (Geritz et al. 1998; McGill and Brown 2007). While this may be true when we are interested in the adaptation of

quantitative traits, we claim that this hypothesis is not appropriate for the evolution of more qualitative aspects of cooperative behaviour. In particular here we take the example of collective hunting. Because the evolution of coordination is necessary for the emergence of the collective behaviour, this means that the mechanistic constraints may have a crucial influence on the emergence of cooperators and the evolution of cooperation. In consequence, we want to show that by considering the mechanics of behaviour as a black box, critical effects are often neglected and that proximate mechanisms influence ultimate explanations.

1.2 Model and Method

Now that we have properly introduced the global question asked in this thesis, we will present the method with which our study is conducted.

1.2.1 Game Theory and the Stag Hunt

It is classic to use abstract models to study the evolution of cooperation, as we will explain more thoroughly in Chapter 2 where we will provide a more extensive review of models used to that end. Models are convenient because they consider general mechanisms and capture the relations between key factors. From purely computational models to spatial simulations, the toolbox of models has been expanded during the past decades in order to increase our understanding of the evolution of cooperation.

Among all these types of models, the most famous for studying cooperation dilemmas are game theoretical models. The principle is that each game represents a particular social interaction between several (most often two) players. Each game is defined by a *payoff matrix* whose goal is to indicate for each player, given her strategy and that of the other player, what is her expected reward. This way, the payoff matrix is used to describe the specificity of the game as a whole. Game theoretical models are well used in economics and some of them, like the *Prisoner's Dilemma*, are even highly popular outside of the scientific community. As such, evolutionary biologists have also been interested in using game theory as a way to study the evolution of cooperative behaviours. In the case of evolutionary biology, the general framework is called evolutionary game theory (Maynard Smith and Price 1973).

In the context of this thesis, we are interested in mutually beneficial actions that require coordination between several individuals. As such, we focus on a particular type of games: coordination games. The most well-known representent of coordination games is called the *Stag Hunt* (Skyrms 2004). Following a metaphor of the social contract introduced by Jean-Jacques Rousseau and then popularized as a game by Brian Skyrms, the stag hunt follows a simple story (see Figure 1.2)). Two hunters have the choice of either hunting a hare or a stag. Catching a hare is easy for any of the hunters and these prey are present in such availability that we can consider that hunting a hare has no influence on the other hunter's strategy. However, a stag represents a much more challenging prey to hunt and hunters need to cooperate if they want to reap the benefits of hunting a stag. Finally,

a stag is more rewarding than a hare. Thus there is a real incentive to cooperate but also a risk that trying to cooperate when the other individual does not will not yield any benefit. The exact payoffs are not what matters most as long as the order between each situation is respected. In the case of the stag hunt, the order for each hunter is as follows: successful cooperation on stag, hunting hare (whatever the partner's strategy) and failed attempt at cooperation.

		Hunter 2	
		Hare	Stag
Hunter 1	Hare	P 1 ; 1	T 1 ; 0
	Stag	S 0 ; 1	R 2 ; 2

	Payoff-dominant Equilibrium
	Risk-dominant Equilibrium

Figure 1.2 – Payoff matrix of the stag hunt. In the stag hunt (Skyrms 2004), we consider that while hunting, two hunters can either hunt a hare or a stag. Hunting a hare can be done in a solitary or cooperative fashion, which ensures that any individual which hunts gets a reward. In comparison, hunting a stag can only be achieved in a cooperative fashion but rewards more than a hare. In consequence, an individual who would hunt a stag alone would not get any benefit. Payoffs are indicated in pair as follows: (Payoff for hunter 1; Payoff for hunter 2). The exact payoffs do not represent the most important aspect of the game as long as the different situations are in that order: R (reward for cooperation) $>$ T (temptation for defection) $= P$ (punishment for defection) $>$ S (sucker's payoff). The payoff-dominant equilibrium is the equilibrium where the hunters maximize their maximum payoff whereas the risk-dominant equilibrium is the one where they maximize their minimum payoff.

The particularity of this game is that there are two evolutionary stable Nash equilibria (Nash 1950; Maynard Smith and Price 1973). This means that when both individuals hunt hare or when both individuals hunt stag, either strategy is stable against the invasion of mutants. More precisely, this implies that in comparison to the prisoner's dilemma, when the cooperative equilibrium is evolved (hunting stag), its stability is not threatened by the invasion of "defectors" (hare hunters). Thus, in coordination games, cooperation is stable once evolved but risky when rare (Forber and Smead 2015). The stag hunt is thus appropriate to study the emergence of cooperative actions when faced with a bootstrapping problem (i.e. the spread of initially rare cooperators).

But as previously said, when we abstract from the mechanics of behaviours as in game theory we may neglect the critical influence of these proximate aspects on the evolution of cooperation. When studying the stag hunt we give no explanation on the origin of the coordination mechanisms which allow the individuals to reap the benefits of stag hunting (Calcott 2007). In particular, it is often assumed in game theory that a single mutation is sufficient to switch from one equilibrium to the other. In reality, cooperation

cannot be beneficial unless coordination is evolved and coordination is not beneficial on its own. In consequence, the emergence of the cooperative equilibrium entails complex modifications in the behaviours of individuals. Thus the mechanics of these behaviours may impact the evolution of this equilibrium.

Our aim is thus to model the practical mechanics of coordination behaviours to study how they influence the emergence of cooperation. To this end, our approach is that of modeling in *evolutionary robotics*² (Nolfi and Floreano 2000).

1.2.2 Evolutionary Robotics

Evolutionary robotics (ER) is a method based on designing robots by taking a loose inspiration from natural evolution. Namely, ER takes the concepts of *selection* and *variation* in order to explore the complex space of candidate solutions for the design of a whole robotic system. The idea of using evolutionary processes in order to solve engineering problems is not new. The whole field of evolutionary computing was created on this idea and offered success in optimization problems where more classical methods fail (Holland 1975; Goldberg 1989; Eiben and Smith 2015). Evolutionary robotics use the same principles to take on the complex task of designing part or all of a complete robot: sensors, morphology and control (Nolfi and Floreano 2000; Floreano et al. 2008; Doncieux et al. 2015). Keep in mind that the term "robot" is used loosely here and can refer either to a physical or simulated robot. This does not impact the general method of evolutionary robotics.

Evolutionary robotics is constituted of an evolutionary algorithm whose goal is to evolve a population of artificial genotypes according to a fitness function (see Figure 1.3). While the actual format of the genotypes is of no particular interest here, it is but rarely similar to a real genotype in both complexity and features. One classical choice in ER is to use a collection of real values for each genotype but some other popular choices are to use booleans (i.e. genetic algorithms) or data trees (i.e. genetic programming) (Eiben and Smith 2015). This genotype is then translated into a phenotype which constitutes the robot's morphology and/or control. Again, the transition from genotype to phenotype as well as the actual phenotype itself can both vary greatly from one model to another. On that matter, one must choose what best fits his/her needs. The important point is that in ER the phenotype is what is evaluated. To that end, the robot is situated in its environment and let to interact with the environment and/or other robots.

In the more classical models, a fitness function is used to compute the fitness score based on the behaviour of the robot in its environment. For example, in one of the first experiments where an evolutionary process was used to automatically design the control of a robot (Floreano and Mondada 1994), the goal was for a robot to navigate a looping maze. As such, the fitness function that had to be maximized was designed as follows:

²The choice of using evolutionary robotics is not without consequences and is not made arbitrarily. There are critical reasons which justify that we use this technique rather than any other among the numerous modeling frameworks available in evolutionary biology. As this first Chapter is a general introduction to this manuscript, we will carefully motivate this choice in Chapter 2. In addition, a brief motivation will be provided in the next Section.

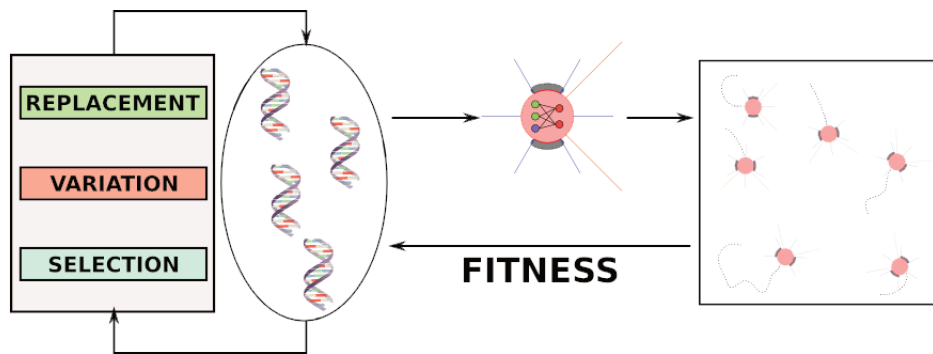


Figure 1.3 – General workflow of an evolutionary robotics algorithm. The main goal of ER is to evolve a population of genotypes. To that end, each genotype must be evaluated to obtain a fitness score. A genotype is thus translated into a phenotype (here an artificial neural network) and then embedded into a robot to act as its controller. The robot is situated in its environment and its behaviour is evaluated in accordance to the specificities of the task. Once every genotype has been assigned a fitness score, they undergo an evolutionary algorithm. This process selects the genotypes deemed fit to create offspring, on which variation is then applied. Finally this new population of genotypes replace the previous population and the process can go on for a new generation.

$$F = V(1 - \sqrt{\delta v})(1 - i)$$

where (1) V is the average rotation speed of the wheels, (2) δv is the absolute difference between the speeds of the wheels and (3) i the normalized maximum activation value between all sensors (where these sensors were infrared sensors capable of detecting obstacles). Thus, the goal of the robot was to (1) maximize its translational speed, (2) minimize its rotational speed and (3) minimize the activation of its sensors. In other words its goal was to move (1) as fast as possible (2) as straight as possible and (3) by avoiding obstacles at the same time. A large part of designing an evolutionary robotics algorithm may be spent carefully crafting the fitness function to evolve the behaviour of the robot according to what is desired (which is a challenge on its own). However a recent shift occurred in the field toward algorithms that do not rely on a complex and handcrafted fitness function. Rather there is a strong interest in methods that go beyond the focus on objective functions to concentrate on the search for novelty (Lehman and Stanley 2008, 2011) or diversity (Mouret and Doncieux 2012) in the behaviours evolved. Lastly, it is interesting to note that in most instances of evolutionary algorithm, fitness is thus used to guide the evolutionary process. This is obviously different from the biological definition of fitness which is an *a posteriori* observation of the capacity to produce offspring.

After genotypes have been evaluated, there needs to be the insertion of new individuals in the population. This can be done in a generational fashion, where most or all of the population is replaced at each generation or in a steady state manner, where only a few selected individuals are removed from the population and replaced by new individuals. In order to decide which genotypes will be able to produce offspring and create these new

individuals, a selection scheme based on the fitness score is applied. Some of the most popular methods for the selection of individuals are $(\mu + \lambda)$ -ES, *fitness proportionate* and *tournament-based* selection. The first two methods will be thoroughly described in the next Section. As an example, we quickly describe the process of tournament-based selection. Under this scheme, an offspring results from a tournament between several (from two to population size minus one) randomly chosen genotypes in the population. These genotypes are then ranked by fitness score. Next the "winning" genotype is selected based on a given parameter T . A random value p is generated. If $p < T$ then the best genotype is selected. Otherwise we select one of the other genotypes with the same method. This parameter is used to tackle the tradeoff between exploration and exploitation. It is often absent from tournament-based selection as the size of the tournament can also address this dilemma.

Variation is then finally applied on the offspring to create the population of the next generation. Variation can consist of mutations and/or crossover. A mutation is the process of randomly choosing one or several genes in the genotype, for example according to a uniform distribution, whose value is then randomly changed (in the way that depends from the format of the genotype). In comparison, crossover is used to mix the genotypes of two different offspring. In the most classical way to do crossover, one point crossover, a random point is selected in the genotype and genetic material is swapped between two individuals around this point. Numerous operators of variation exist and which one to use depends on the problem at hand.

In this thesis, we are interested in the modeling of proximate mechanisms in the evolution of mutually beneficial actions. As such, evolutionary robotics is a suitable approach. ER focuses on the modeling of individual-level behaviours resulting from the evolved genotypes situated in a specific environment. This allows to take into account interactions with both the environment and other individuals under complex ecological features. As such, ER has been used to address specific biological hypothesis with a strong emphasis on the mechanistic constraints at play in these evolutionary phenomena (Floreano and Keller 2010; Mitri et al. 2013). Several recent works demonstrated the convenience of ER in this context by investigating for example the evolution of cooperation (Waibel et al. 2011, 2009), the evolution of communication (Mitri et al. 2011; Wischmann et al. 2012) or division of labour (Ferrante et al. 2015). These works will be carefully reviewed in Chapter 2, where we will also provide a more detailed justification of our modeling choice. In consequence, ER is a fitting choice to model the proximate mechanisms of coordination and study their impact on the evolution of cooperation: the mechanics of behaviours are not considered a black box anymore.

Around the common theme of the evolution of coordination, we also want to address an additional problem which is the design of cooperative robots. This is another manner in which evolutionary robotics can contribute to scientific research (Trianni 2014; Doncieux et al. 2015). In particular, ER is a valuable approach when designing multirobot systems. The design of multirobot systems is complex because it requires to take into account the interactions between multiple individuals in order to produce the emergence of collective functionalities. The automatic design of robot control with classical learning methods in particular is challenging because of the sheer complexity produced by the size of the

problem. In comparison, ER can be used as a black-box optimization technique which does not need to make approximations about the problem at hand. In consequence, several works have focused on designing multirobot systems with ER, be it for the design of robot swarms (Baldassarre et al. 2007), the evolution of specialisation (Ferrante et al. 2015) or the design of flying communicative robots (Hauert et al. 2014). In this manuscript we thus focus on the automatic design of multirobot systems in evolutionary robotics. In particular, we study the influence of team composition (i.e. homogeneous or heterogeneous individuals) on the evolution of cooperative robots (Waibel et al. 2009). We will introduce this problem more thoroughly in Chapter 5.

1.2.3 Experimental Setting

We are interested in the dual objective of both modeling the evolution of cooperation and designing multirobot systems with an evolutionary robotics approach. Both these aspects are independent and may be considered separate problems. However these two facets share the common problem of the nature of coordination behaviours in the evolution of cooperation and as such we use a similar setting for each of them. In particular, our inspiration is the framework of the stag hunt which we model in evolutionary robotics. As previously explained, this gives us the possibility to model the mechanistic constraints at play in the evolution of mutually beneficial cooperation. However it also serves as an appropriate inspiration for designing multirobot systems. In particular, we want to investigate the nature of coordination behaviours between heterogeneous robots (in terms of robot control) and the influence of such team composition on the evolution of cooperation when selfish behaviours are possible. In consequence, while both our problems are separate we chose to use a similar (or at least strongly similar) experimental setting for each of these approaches. In this Section, we present this setting. As they may change depending on the exact experiments presented in this manuscript, some of the parameter values are not specified in this Section (see Table 1.1). All the parameters that we use have been chosen after conducting preliminary studies.

Robot model

We want to study the evolution of simulated robotic agents (see Figure 1.4). These agents are capable of movement thanks to two independent wheels and are equipped with a collection of sensors. Those sensors are of two types: 12 *proximity sensors* and a 90 degrees front *camera*. On the one hand, the proximity sensors are equally distributed all around the robot's body and return the proximity of any obstacle nearby (i.e. in a radius which equals twice of the body's diameter). On the other hand, the camera cannot sense obstacles but can feed the agent with the type of any object its rays collide with in the environment (including other agents). More precisely, this camera is composed of 12 rays with an infinite range equally divided in the camera's angle. When one of these rays intersects with an object, it returns the type of this object and its proximity. The robot is thus constituted of simple sensory capabilities. The choice of having two different sensory feedbacks is not innocent. By dividing the sensory capabilities between the proximity

Parameter		Value
Evolutionary Algorithm		
	Per locus mutation probability	5×10^{-3}
	Mutation operator	Gaussian $\mathcal{N}(0, 0.01)$
	Number of partners	5
	Number of simulations per pair	5
Artificial Neural Network		
	Input neurons	49
	Hidden neurons	8
	Output neurons	2
Simulation		
	Simulation duration (in time steps)	20000
	Capture duration (in time steps)	800

Table 1.1 – *Experimental parameters.*

sensors and the camera, we are essentially facilitating the process of evolving two basic skills necessary for the robot: obstacles avoidance and agents recognition. This design is not to be considered as a realistic approach to animal modeling but rather as a way to ease the acquisition of these skills that are of no particular interest here. Furthermore, while the obstacles avoidance mechanism is not expected to improve much during evolution, the appearance of cooperative behaviours in comparison should lead to variation on the manner with which to recognize agents.

Controller

The controller of each agent is an artificial neural network (ANN). While a lot of different types of controllers are used in evolutionary robotics, ANN are widely employed for their versatility (Doncieux et al. 2015). The principle behind a very basic neural network is that it is constituted of a layer of input neurons and a layer of output neurons which are connected (sometimes fully) to each other. Each of the connection has a value, which is called a connection weight. The value of each output neuron is computed as the sum of the input neurons connected to it weighted by the connection weight. A transfer function can then be applied to this output to compute the final value. ANN are really diverse in how they are implemented and can include recurrence or have their topology evolve (Stanley and Miikkulainen 2002).

In our case, we use a fully connected multilayer perceptron with one hidden layer. This neural network is composed of two outputs which are used to compute the speed of each of the robot’s wheels. The inputs of the network are constituted of all the sensory information of the robot in addition to a bias neuron whose value is always 1. This amounts to a total number of 49 input neurons: 1 for each proximity sensors, 3 for each camera ray and 1 for bias. The feedback from each camera is encoded by 3 neurons as we use 2 bits to encode the type of each object (hare, stag or the other agent) and 1 last

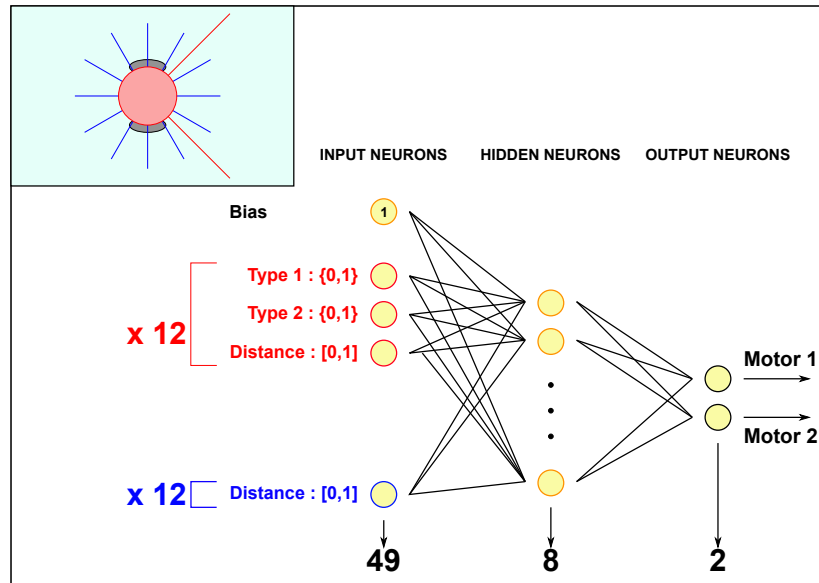


Figure 1.4 – Robot model of our experimental setting. This figure represents the sensory and neural architecture of the simulated robotic agents used in our experimental study. On the robot diagram, proximity sensors are represented by the blue lines whereas the front camera is shown as a red cone. The neural network is a multilayer perceptron with one hidden layer and whose inputs are constituted of all the sensory information of the individual. The outputs of the neural networks are the speeds of both of the robot’s wheels.

neuron for the proximity of this object. The hidden layer is constituted of 8 neurons. Finally, the transfer function used in each neuron is a sigmoid and the topology of the ANN is never changed throughout evolution.

Environment

We place two evolved robots in an arena with four solid walls. This arena is filled with randomly positioned objects of different types, where the type can be recognized by the camera. These objects represent the prey that can be hunted by the individuals in our modeling of the stag hunt game. The objects cannot move while the robots can move freely. We conducted preliminary experiments with moving objects but it was shown that this did not significantly impact the behaviours of the individuals. Additionally, the presence of multiple stationary objects already implies that the individuals need to evolve coordination. In consequence, the addition of motion for the objects is not critical for the conduct of our study.

In order to catch an object, an individual needs to move to this object and then stay next to it for a specified amount of time steps (800 time steps out of 20000). After this duration, the object is removed from its position and replaced at another random position in the environment; we thus ensure a constant ratio of each type of object. For cooperation to occur, both robots need to be close to the object at the end of this duration. This thus

implies that robots need to display actual coordination behaviours in order to be able to cooperate. This also means that an individual can reap the benefits of cooperation simply by being there at the very last step of the capture period.

An object is always removed if an individual is next to it after this period of time, regardless of whether it requires cooperation. What varies are the rewards given to the individuals. Even in the case of cooperation, we do not study the way rewards are distributed between individuals: they are both equally rewarded.

Evolutionary algorithm

The genotype of each individual is constituted of all of the connection weights of the neural network. Each gene is initially randomized in the interval where it takes its values, i.e. in $[0, 1]$. To evolve these genotypes we use a classical evolutionary algorithm. At each generation of the algorithm we evaluate each individual from the population in the arena presented before. Its partner is randomly selected in the population. To ensure that each individual encounters a fair sample of the population, each individual is separately paired with 5 different partners. From a biological perspective, this means that encounters between individuals are rare w.r.t. the life of a given individual. Then a pair of individuals interacts in the arena during 20000 time steps. In order to decrease the effects of stochasticity due to the objects' random positioning, each pair plays 5 different simulations. Thus, each individual plays a total number of 25 simulations. Fitness is obtained by computing the average reward of the individual in these simulations.

The individuals are then selected to produce offspring. Throughout our experiments, we mainly study two different selection methods: *fitness proportionate* and *elitist*. The former is the more classical one when modeling evolutionary biology because it corresponds to a Wright-Fisher model (Wright 1931) with constant population size. Under this model, we randomly sample through the population to select a parent in order to create each offspring that will constitute the population of the next generation. Each individual in the population has a higher probability to be selected if its fitness is higher and can be selected multiple times. The latter selection scheme is implemented as a $(\mu + \lambda)$ -ES. With this selection method, we always keep the μ best individuals of each generation for the next generation. Then we add λ offspring to the population of the next generation, where the parents of these offspring are taken from the μ best individuals ranked by fitness score. In the case of biological modeling, fitness-proportionate is thus a more realistic choice. However, we observed that the elitist selection strategy would reach similar results as fitness proportionate but with smaller population sizes. Thus it allowed to decrease the computational time of our experiments.

Whatever the selection strategy, we always create the offspring in the same way. Each offspring is a mutated clone of its parent. Then mutation is applied independently on each gene according to a mutation rate of 5×10^{-3} . If a gene mutates, mutation is sampled according to a gaussian distribution $\mathcal{N}(0, 0.01)$. Lastly, we use no recombination (i.e. crossover) in any of our experiments. Because this thesis is focused on cooperation between unrelated individuals (in model and in design), it raises no particular issue to limit variation to mutations.

1.3 Evolving Coordination in Evolutionary Robotics

While this introduction was mainly focused on the biological aspect of cooperation and the problem its origin poses for evolutionary biology, we want to study several facets around this general problem. We believe that our approach in evolutionary robotics entails that the contributions of this thesis can serve different purposes around the common subject of the evolution of coordination. Historically, evolutionary robotics has been used at the beginning for the automatic design of robotic systems. However, there has been a debate on how the works in this field could really contribute to scientific research as well as to whom they may be of interest (Trianni 2014; Doncieux et al. 2015). It is now admitted that research in evolutionary robotics should be clearly directed toward either of two goals: modeling biological phenomena or designing robots (Trianni 2014). In this thesis, our goal is to present different contributions which separately aim for each of these goals. In this last Section, we briefly present the structure of the manuscript.

1.3.1 Modeling the Evolution of Coordination

In the first Part of this manuscript, we use evolutionary robotics in order to model the evolution of cooperation. Because we tend to generally ignore or minimize the practical mechanics of behaviour, the role of coordination in the evolution of mutualistic actions is often underestimated. Yet, the proximate mechanisms of coordination may influence the convergence to a cooperative solution. We thus study how the nature of coordination behaviours and the mechanisms that underlie their evolution may impact the evolution of cooperation. The particular issue we address here could be summarized as follows: *what are the proximate mechanisms which hinder or facilitate the evolution of mutualistic cooperation?*

To that end, we will spend some time in the introduction of this Part to motivate our choice regarding evolutionary robotics, something we deliberately skipped in this general introduction. One reason why the proximate causes of coordination have often been overlooked is that classical models used for studying evolutionary problems may not be appropriate for this particular goal. We believe that among the distinct assumptions made by these models, some are critical if we plan to fully understand the evolution of coordination. However, it is important to make clear that we do not pretend our approach to be a more realistic depiction of nature. Rather we claim that, while we still study a theoretical abstraction of cooperative actions, the assumptions behind our model allow us to study particular mechanisms we believe of importance for this issue.

1.3.2 Designing Cooperative Robots

In the second Part of the manuscript, we study the evolution of cooperation in a team of heterogeneous robots. In consequence, we focus on the automatic design of multirobot systems. As we will explain in details in Chapter 5, multirobot systems have now been investigated for a long time for their advantages over single robots. In particular, they may allow to design more efficient and cheaper robotic systems as well as benefit from

the redundancy of multiple robots to design more robust systems. Moreover, it can sometimes be necessary to have several robots acting at the same time to achieve a particular task. The practical applications of such systems are numerous, in particular in environments where humans cannot go and where using a single and generally more complex robot would simply not be reliable enough. As such, multirobot systems could be used for collective manipulation, building or exploration of hazardous environments. For example, cooperative robots could investigate and perform repairs inside nuclear plants after particularly catastrophic incidents.

However designing this sort of systems is challenging. It is one thing to engineer a factory robot which is programmed to perform a very specific and repetitive task in a controlled environment. It is another to design a robot capable of acting in an uncertain environment and able to adapt to the unexpected. And even more complicated when multiple robots must both possess the qualities expected from a single robot and also coordinate in an efficient manner. Multiple techniques for automatic design have been proposed, especially regarding the control of robots. However, when it comes to dealing with changing environment and uncertainty, the "easiest" way is to design a robot that is capable to learn from previous experiences. Among the learning methods used in this context, ER is a promising one when robots are expected to perform in an open and unknown environment. Therefore we explore how ER can be used for the automatic design of collective robots.

We are interested in the nature of the coordination behaviours that could be evolved between heterogeneous individuals. Heterogeneous teams of robots allow for more diverse behaviours to emerge inside a group of individuals. However, while there has always been a clear interest in designing heterogeneity in multirobot systems (Parker 1994, 2008), most research on the evolution of cooperative robots have been focused on homogeneous groups of individuals (Waibel et al. 2009). This is indeed one of the safest way to ensure that robots will evolve a cooperative behaviour, as there is no selfish interest to act in a solitary fashion. However, we believe that the influence of heterogeneity on the quality of the coordination behaviours is as much of importance as the capacity to evolve a cooperation solution. In particular, when coordination is needed heterogeneity may lead to more efficient cooperative behaviours. However, conflict can arise from the selfish interests of the individuals. Thus, the issue we focus on in this second Part of the manuscript is: *how can we evolve efficient coordination behaviours in a group of heterogeneous robots ?*

1.3.3 Organization of the Manuscript

This manuscript is composed of two Parts that each addresses one of the two problems we described here. Each Part is constituted of an introduction Chapter and two results Chapters. Each of the results Chapters begins with a short introduction and quick summary of the results presented in the Chapter. After the two Parts, a final concluding Chapter is present. Chapter 2 presents a brief overview of the modeling techniques used in evolutionary biology and for the evolution of cooperation in particular. The goal is to provide the reader with a motivation about our choice to use evolutionary robotics to model the evolution of cooperation. Chapter 3 is presented as an article published in an interna-

tional journal. It focuses on a comparison between a classical game theoretical analysis of the stag hunt and the results we obtained with our evolutionary robotics model. Chapter 4 is shown as the draft for a future journal article. This Chapter addresses the issue of optimizing collective actions by way of individual selection and the impact of the nature of coordination strategies in that context. Chapter 5 briefly presents the field of multirobot systems and then review the different methods used to design the control of distributed robots. The design problem of evolving cooperation between heterogeneous robots is then described. Chapter 6 reads as an article published in an international conference. It focuses on a comparison between clonal and aclonal approaches in a cooperative foraging task on evolvability and efficiency. Chapter 7 is presented as another article published in an international conference. It deals with the evolution of specialisation at the level of the population, i.e. the evolution of genotypic polymorphism. Finally, Chapter 8 summarizes our contributions and addresses their limits as well as the perspectives this thesis opens for future research.

Part I

Modeling the Mechanics of Coordination in the Evolution of Cooperation

Chapter 2

Models in Evolutionary Biology and Evolutionary Robotics

So far, we have been able to study only one evolving system and we cannot wait for interstellar flight to provide us with a second. If we want to discover generalizations about evolving systems, we will have to look at artificial ones.

— John Maynard Smith

2.1	Classical Models in Evolutionary Biology	42
2.1.1	Population Genetics	42
2.1.2	Evolutionary Game Theory	43
2.2	Individual-Based Modeling	46
2.3	Evolutionary Robotics	48
2.3.1	Individual-Based Modeling and Evolutionary Robotics	48
2.3.2	The Evolution of Cooperation in Evolutionary Robotics	50
2.4	Conclusions	51

In the first Part of this manuscript, our goal is to model the proximate mechanisms in the evolution of coordination and study their impact on the emergence of cooperation. As in most fields of science, modeling is a standard approach to study questions that cannot be understood by simply observing the physical world. Everything that we now contemplate is the consequence of thousands of millions of years of evolution; the earliest appearance of life is dated back to 4000 millions years ago and a little more than 2000 millions for the eukaryotes. This means that we are mostly left with evidence from the past (i.e. paleontological records) (Aiello and Wheeler 1995; Wrangham et al. 1999) or direct observation of evolved behaviours (i.e. ethology). Evolution can also be studied as it is taking place in organisms where it is a much faster process, like microorganisms¹ (Elena

¹On that subject, a noteworthy experiment is that of Richard Lenski's "Long-Term Evolution Experiment" on *Escherichia coli* (Fox and Lenski 2015), where he set up in 1988 12 populations of the same strain of *E. coli* and observed their evolution since then.

and Lenski 2003). However, while empirical works can indeed study the proximate mechanisms of cooperation, they do not inform on the ultimate explanations of behaviours and are not sufficient on their own to garner a full understanding of the process. Thus models are now commonly accepted in the field of evolutionary biology, even if this may not have always been the case (Shou et al. 2015).

As said in the Introduction of this manuscript, we chose evolutionary robotics as our modeling technique for the problem studied. However, we only briefly justified this choice. Thus, the next Sections will be devoted to this task. We will first focus on the different modeling approaches used in the field of evolutionary biology. To that end, we will distinguish between the more classical methods for modeling evolution and the computational methods that arose with the availability of computers. Rather than doing an extensive review of models, we will present the relevance and benefits of these models with regard to our subject. This will finally give us the opportunity to motivate our approach in evolutionary robotics in light of the set of models available and the problem studied in this thesis.

2.1 Classical Models in Evolutionary Biology

It is nowadays classical in evolutionary biology to use modeling to address the evolution of cooperation (among other evolutionary traits). It is also necessary to do so to fully grasp the mechanisms at play. Models are thus numerous and a lot of different frameworks have been proposed to classify cooperative actions (Dugatkin 2002; Sachs et al. 2004; Lehmann and Keller 2006). To that end, mathematical models dominate the field (Servedio et al. 2014).

2.1.1 Population Genetics

Population genetics form a large part of the literature in evolutionary biology. The inception of this field was mostly the result of Ronald Fisher, J.B.S. Haldane and Sewal Wright. Fisher was the first to link mendelian inheritance (i.e. the inheritance of biological traits) with mathematical models of natural selection in his book *The Genetical Theory of Natural Selection* (Fisher 1930). Population genetics is concerned with studying the changes in alleles' frequencies at a particular locus in the genotype. In particular, the focus is put on population wide variations of evolutionary traits within one or a few loci. As such, there has been a strong emphasis on studying genetic mechanisms like dominance, epistasis or genetic drift. A notable branch of this field, *quantitative genetics* have in comparison been more focused on the phenotypical aspects of evolution. More precisely, quantitative genetics deals with continuously varying phenotypical traits. It thus abstracts from the genetic details of evolution. However, population and quantitative genetics both tend to abstract from the role of ecological features which we are interested in here.

2.1.2 Evolutionary Game Theory

Matrix games

In comparison to population genetics, *evolutionary game theory* (EGT) puts a strong emphasis on ecological aspects. Game theory was originally conceived by the mathematician John von Neumann as a way to determine the optimal strategies in a contest between several (usually two) "players". Given a payoff matrix, each player can expect a certain payoff depending on his strategy and that of other players. In this framework, players are expected to be rational and follow this optimal strategy. One of the most important concepts of game theory is the *Nash equilibrium* (Nash 1950). Under a Nash equilibrium, no player benefits in deviating from his strategy if the other players keep their strategy. This framework was first adapted to darwinian evolution by John Maynard Smith and George Price under the name of evolutionary game theory (Maynard Smith and Price 1973). The novel idea is that the players' strategies are based on their phenotypes rather than a rational choice. Therefore, an individual's strategy is now inherited and the evolutionary success of the different strategies are studied. To that end, the payoff matrix of an evolutionary game corresponds to the fitness value of each strategy. In order to study the evolutionary dynamics of an evolutionary game, we consider a population of individuals all playing the same strategy. We then assume the appearance of a rare mutant who plays a different strategy and study the fate of these two strategies. If the strategy of the mutant has a higher fitness than that of the initial population (called the *resident strategy*), then it may invade the population and replace the resident strategy. Otherwise, the mutant strategy will not be favoured by selection and will disappear. If this resident strategy is stable against any mutant strategy, we say that this strategy is evolutionarily stable (ESS) (Maynard Smith and Price 1973). Interestingly, all ESS are Nash equilibria (but the opposite may not hold).

One of the main feature of evolutionary game theory in comparison to population genetics is that it takes into account the influence of an individual's behaviour on the fitness of others. More precisely, the fitness of an individual depends on the proportion and behaviours of other individuals in the population, which is known as *frequency-dependent selection*. As such, EGT is convenient in order to account for the ecological features of a particular evolutionary phenomenon (Hammerstein and Selten 1994). Because it represents conflictual and cooperative interactions, there is a great interest in using EGT for the study of the evolution of social behaviours (Bshary and Oliveira 2015). A large body of work in particular has focused on the stability of altruism when faced with the appearance of free-riders (defectors) in the prisoner's dilemma (Requejo-Martinez 2013). The prisoner's dilemma is a famous game where the only evolutionary stable strategy is for both individuals to defect. Therefore studying how the cooperative equilibrium could be stable in this situation is challenging. Another famous example is that of reciprocity in the *Iterated Prisoner's Dilemma* (Axelrod 1984). Axelrod & Hamilton proposed that individuals play an iterated version of the prisoner's dilemma (of which you can find Axelrod & Hamilton's payoff matrix in Table 2.1). To put it more simply, when individuals have engaged in one prisoner's dilemma interaction, there is a probability that they will meet

again in a later interaction. They showed that, under those circumstances a particularly efficient strategy was one called *Tit for Tat* (TFT). Under this strategy an individual always cooperates when meeting an opponent for the first time. It then always copies the opponent's last move which means that it (1) retaliates and (2) does not hold grudges. This strategy was thus presented as a theoretical example of reciprocity (Trivers 1971).

	Cooperation	Defection
Cooperation	3,3	0,5
Defection	5,0	1,1

Table 2.1 – Payoff matrix of the prisoner's dilemma. The strategy of player A (resp. player B) is symbolized by each row (resp. column). The payoff of player A (resp. player B) is shown on the left (resp. right). This is the payoff matrix which was used by Axelrod & Hamilton in their work on *Tit for Tat* in the prisoner's dilemma (Axelrod 1984).

In the case of coordination which we are interested in here, the prisoner's dilemma is not appropriate to model the evolutionary dynamics of cooperation (Alvard and Nolin 2002; Skyrms 2004). As such, another type of games more suitable for this question was introduced as coordination games and in particular the stag hunt (Skyrms 2004; Requejo-Martinez 2013). The details of this game have already been covered in the Introduction. The major difference with the prisoner's dilemma is the presence of a second ESS as the cooperative equilibrium (i.e. stag hunting). Thus the emphasis of this game is not on the stability of a population of cooperators against the invasion of free-riders (as the cooperative equilibrium is evolutionarily stable) but on the transition from the solitary equilibrium to the cooperative one. In his book, Skyrms mainly studied the influence of location, signaling and partner choice on the evolution of cooperators (Skyrms 2004). However, coordination games like the stag hunt have received relatively little attention, especially in comparison to the prisoner's dilemma (Iyer and Killingback 2016).

Adaptive dynamics

In classical EGT (i.e. matrix games), strategies are discrete and generally constitute a finite list. In comparison, most evolutionary traits take values in a continuous domain. We can for example think about the size, flowering rate or investment and allocation of resources (McGill and Brown 2007). In order to study the evolutionary dynamics of those traits, a continuous version of EGT rose as *adaptive dynamics* (or continuous-trait game theory). This modeling technique can be seen as a way to combine quantitative genetics, by studying the rate of change of a population's strategy, and EGT, by applying the ecological aspect of frequency-dependent selection (Geritz et al. 1998; McGill and Brown 2007). More precisely, adaptive dynamics extends on the notion of evolutionarily stable strategies from EGT. In particular, the concept of ESS as it exists in EGT lacks precise knowledge about the convergence of a given strategy. Namely, we know that an ESS may not be invaded by any mutant strategy once it has spread in the population. Yet, we do not know if this strategy will eventually become established given the initial conditions. This problem of convergence is represented by *convergence stability*. Convergence stability

implies that a strategy, thanks to multiple small evolutionary steps, will be able to fix in the population. Both concepts of evolutionary stability and convergence stability do not always come together (Eshel and Motro 1981; Eshel 1983). Behind convergence stability is the idea that the shape of the fitness landscape changes as the resident strategy changes. From this it stems that it may be impossible to evolve an ESS even when it represents a fitness maximum.

Several key evolutionary concepts that could not be modeled by classical EGT have been introduced through the framework of adaptive dynamics. One such concept is that of "branching points". These occur when a strategy is convergence stable but not evolutionarily stable (Geritz et al. 1998). Namely this strategy acts as an evolutionary attractor from afar but, because the fitness landscape changes as the resident strategy changes, this strategy may be a fitness minimum (and thus not ESS). These are called branching points because two different evolving populations may coexist and evolve separately. Branching points have thus been used to model the evolution of speciation (Geritz et al. 2004).

Proximate mechanisms in evolutionary game theory

Models in EGT represent a classic framework for the study of the evolution of cooperation. However, they make assumptions that we deem critical for the evolution of mutualistic cooperation. In matrix games, it is often considered that any given strategy can evolve regardless of the resident strategy. In particular, the phenotype of a given individual is often simply modeled as either a "Cooperator" or a "Defector". As such a single mutation is sufficient to evolve one phenotype or the other. In the context of adaptive dynamics, the issue of converging towards a strategy under particular ecological context (i.e. convergence stability) is well studied. However, convergence is addressed under the assumption that the effects of mutations are small and that convergence can be achieved through a series of small evolutionary steps.

In both cases, these models make assumptions on the availability of mutations; they are considered not to be limiting and the mapping between the genotype and the phenotype is thus not explicitly modeled. As we already explained, the effect of such assumptions is not equivalent depending on whether we are interested in the stability or the origin of cooperation. If we find a strategy to be an ESS when mutations are not limiting (i.e. any mutant can appear in the population), then this strategy is also an ESS under stronger limitations. In comparison, the evolution of such strategy may vary depending on the effects of mutations. In particular, there needs to be a succession of mutants which are each favored by selection between the different equilibria. In consequence, the availability of mutants may be critical for the origin of cooperative traits. Because mutations affect the proximate mechanisms of behaviours, then the nature of these mechanisms is crucial to the appearance of cooperative mutants. In particular, this may be true when the ability to coordinate is necessary for cooperation to be beneficial. Therefore, we believe that there needs to be complementary models that consider these mechanics of behaviour in the context of the evolution of cooperation. Namely, we are interested in models where the individuals are explicitly modeled and thus where minimal assumptions are made on the mechanistic constraints at play in the mapping from genotype to phenotype.

2.2 Individual-Based Modeling

The mathematical models presented before are oftentimes labelled as "classical models" (DeAngelis and Mooij 2005; Adami et al. 2014) and this is how we will refer to them in this manuscript. This term is not used in a derogatory fashion. Rather, this is a way to discriminate between the mostly mathematical models which have been classical in evolutionary biology and a range of models which were born thanks to an easier access to computational power. This allowed to approach biological questions from a different direction that, some would argue, enables to go beyond what is possible with purely mathematical models (Adami 2012). However, the line between classical and computational models can sometimes be not so easy to draw and there is a real scientific interest in trying to get the best of both worlds (Wilson 1998). In our case, we are interested in the modeling of individuals so that we can take into account the mechanistic constraints in the evolution of coordination behaviours. To that end, we now present the field of *individual-based models* (IBM)² (Huston et al. 1988).

As stated in the name, the goal of an IBM is to model individual-level mechanisms. This is very different from most classical models in ecology where the emphasis is mainly put on population dynamics (Grimm and Railsback 2005). This does not mean that research in IBMs do not deal with population dynamics but rather that these dynamics are studied as a consequence of the interactions between individuals. The main focus of an IBM is to study the collective dynamics emerging from individual-level interactions (whether with other individuals or the environment). And more importantly, the particularity of an IBM is that these individuals, which are the building blocks of the system, are the results of adaptation: collective properties arise from these (sometimes simple) individual behaviours.

IBMs are mostly used to study biological phenomena for which individual variations, and the assumptions that stem from them, are important. DeAngelis & Mooij (DeAngelis and Mooij 2005) set forth five axes along which IBMs are used to model mechanistic details in the variations between individuals:

Spatial variability While classical models sometimes take into account spatial organization, IBMs allow to model local heterogeneity between individuals.

Life cycle The variability of ontogenetic history can be modeled with finer details by using IBMs than with classical models.

Phenotypical variation and plasticity The influence of individual experience on behaviours can be taken into account. In particular, IBMs can be more appropriate to model the interactions of multiple different behaviours than classical (game theoretical) models.

²The term *agent-based model* can often be found in the literature in lieu of individual-based model. Both names refer to the same framework and are interchangeable. While individual-based models are found more often in biological applications (Grimm and Railsback 2005), no real consensus exists on which term to use. We choose to use the latter throughout this manuscript.

Learning Learning is a consequence of lifetime interactions which are dependent on individual variations.

Genetics and evolution The computational power of IBMs can help study complex evolutionary genetics.

Thus IBMs have been widely used in behavioural ecology for very diverse applications (DeAngelis and Mooij 2005). For example, spatial variability has been of great interest for the study of group patterns. Most notably, models on the formation of groups of animals, whether swarms of insects, flocks of birds, herds of mammals or schools of fishes (Huth and Wissel 1992; Reynolds 1992; Gueron et al. 1996; Couzin et al. 2002) rely heavily on the framework of IBMs. Aggregation behaviours were found to easily arise from simple local (individual-level) sets of rules, leading to complex collective behaviours. More generally, IBMs have been used to model ecological phenomena as diverse as the optimal gap between trees in forests models (Botkin et al. 1972), movement patterns in prey-predators interactions (Smith 1991) or differences in foraging between solitary birds and large flocks (Toquenaga et al. 1995).

Interplays between IBM and EGT are numerous. Indeed, both methods can be used to focus on the evolution of phenotypes and rely heavily on the ecological features of the system. This led to a great number of research bringing together these two fields "with ease". In particular, it is now common to take spatial interactions into account (e.g. how individuals are located on a graph or network) when studying the evolution of (mostly altruistic) cooperation (Hauert and Doebeli 2004). IBMs also give the possibility to more accurately predict the effects of a finite population size while most models in EGT use infinite population sizes (Hauert et al. 2009). Additionally, it is possible to more easily model stochastic or conditional strategies. These ecological features can be modeled with classical game theoretical models but at the cost of increased mathematical complexity (Hauert et al. 2009).

Moreover, individual-level modeling implies that minimal assumptions are made about the effects of mutations on the evolution of individual behaviours. This makes IBMs an interesting addition to EGT when dealing with the modeling of proximate mechanisms. By putting the emphasis on the individual, it is thus possible to clearly study how individual adaptation can lead to the fixation of evolutionary traits at population level. In consequence, some have used IBMs to study the evolution of cooperation. Olson and colleagues have been interested in the evolution of herds (Olson et al. 2013b). More precisely, while collective aggregation benefits the individuals in the group, it is also costly for them (i.e. sharing resources and increasing the risk of being spotted by predators). Thus there is an evolutionary question on the origin of such collective behaviours. They confirmed that the formation of herds could be explained by Hamilton's theory of the selfish herd (i.e. aggregation emerges because every individual tries to put others between itself and the predator) (Hamilton 1971). They also showed that predator confusion through aggregation could explain the evolution of such behaviours (Olson et al. 2013a). Additionally, they revealed that predator confusion may also lead to the coevolution of morphology (vision system) and behaviour for both predators and prey (Olson et al. 2016). Finally they

showed that group vigilance could allow the appearance of gregarious foraging behaviours without any kinship relations (Olson et al. 2014). Others have been interested in symbiogenesis, which refers to the creation of a new species through the symbiosis of previously independent species. Watson et al. showed with an IBM that such process could occur without any relatedness between the individuals (Watson and Pollack 2000). Wilder & Stanley used both an IBM and a classical analytical model to show that altruism could evolve thanks to the creation of ecological niches (Wilder and Stanley 2015).

Lastly, some have been interested in the individual-based modeling of biological phenomena thanks to the simulation of digital organisms. In AVIDA (Lenski et al. 1999), the organisms are programs that compete for CPU time and evolve in a digital environment and the whole framework can be used to model bacterial evolution. The genome of an individual is composed of a sequence of instructions. With this framework, Goldsby and colleagues (Goldsby et al. 2012) modeled the evolution of division of labour (or specialisation), where individuals specialise between different roles. They showed that the evolution of specialists was more frequent when task-switching costs were high. Additionally, they observed that the individuals were able to use stochastic information, location awareness and messaging in order to specialise. The Aevol platform (Knibbe et al. 2005; Batut et al. 2013) also aims at simulating *in silico* bacterial evolution with digital organisms but with an emphasis on a realistic modeling of genomes. As such it has also been used to model the evolution of cooperation. For instance Frénoy et al. (Frénoy et al. 2013) simulated the evolution of cooperation through the production of public goods (i.e. the costly secretion of nutrients). They investigated the influence of genetic architecture for the maintenance of cooperators in the population. They demonstrated the role of second order selection in the protection of cooperation by the entanglement of the genetic architecture.

2.3 Evolutionary Robotics

2.3.1 Individual-Based Modeling and Evolutionary Robotics

In this thesis, we focus on a particular type of individual-based models: *evolutionary robotics* (ER) (Nolfi and Floreano 2000; Doncieux et al. 2015). The technical details of this framework have already been covered in the Introduction of the manuscript. As such we want to present in this Section the reasons for which one could use evolutionary robotics above (or in addition to) any other modeling techniques presented before. Why should evolutionary biologists be interested in using ER ? And why should we use ER in the context of this thesis ?

The main addition of ER when compared to more general IBMs is that ER models are IBMs where the individual is an embodied agent (Mitri et al. 2013). By definition, robots have a body (physical or simulated). In consequence, this creates an additional level of interactions with the environment. Sensory feedback is also part of a robot's design, which means that there often is imperfect information about the environment. In comparison, IBMs will usually (but not necessarily) provide a global and perfect description of the world. Finally, modeling in ER implies that the environment (whether simulated or not)

exists in a bounded space, which again may not be the case in IBMs. All of this can have lasting consequences on the dynamics of a system. When those physical properties are expected to be of importance for the phenomenon studied, it may be beneficial to use ER rather than IBMs. Mitri and colleagues (Mitri et al. 2009) provided an elegant example of a case where physical embodiment led to unexpected results. While their study was focused on the evolution of communication between simulated robots for foraging, they found that the aggregation of robots on a foraging site provided additional information that did not require the use of direct communication. More generally, the line between IBMs and ER in the literature can be blurry. Oftentimes, the differences between both frameworks mainly rest upon terminology and history. Namely, these two techniques come from (at least originally) different communities. ER is deeply rooted in the field of robotics design. In comparison, a large part of the works in IBM is interested in the design of more general multi-agent systems and their applications. While ER and IBM can be very similar, these historical divergences tend to have a lasting effect. Yet, we can simply consider ER models to be a particular instance of IBM as we do here. As such, ER is endowed with similar advantages w.r.t. modeling proximate mechanisms in the evolution of behaviours as IBM.

There has been an extensive, though recent, effort in using ER as a modeling tool for social behaviours as well as evolution (Mitri et al. 2013; Trianni 2014; Eiben 2013; Doncieux et al. 2015). However there still is a lack of communication between communities which implies that ER research sometimes fail to reach those who could be interested by these findings in the evolutionary biology community. This also means that some works in ER may sometimes focus on questions that are of no particular relevance for evolutionary biologists (Trianni 2014; Doncieux et al. 2015). We now present some of the more significant works in ER that have been interested in modeling the evolution of social behaviours.

First, the evolution of communication has been a major subject of interest in ER. The modeling of individual-level interactions is adamant in understanding the evolution of communication. Additionally, the embodiment of individuals in space can have some unexpected effects on communication behaviours (Mitri et al. 2009). Floreano & colleagues (Floreano et al. 2007) showed how the evolution of communication could vary depending on the relatedness inside a group of foraging robots. Robots had to correctly choose between a food site and a poison site, where the difference between both sites could only be determined at close range. Communication could easily evolve when there was strong relatedness (i.e. robots were clones of each other). In comparison, when individuals were unrelated, deceptive strategies would also evolve. In a similar setting Mitri et al. (Mitri et al. 2011) revealed a strong correlation between signal reliability and relatedness between individuals. Related individuals would produce more reliable signals in order to direct others towards the food source. Wischmann, Floreano and Keller (Wischmann et al. 2012) conducted a study where they observed that purely historical contingencies could lead to divergences in the communication strategies evolved in independent evolutionary runs. In particular they showed that signaling strategies of varying complexity could evolve based on these contingencies. The more complex strategy would not ensure higher performance unless in a competitive setting between different populations (a set-

ting in which populations were not evolved). Finally, Mitri and colleagues (Mitri et al. 2009) focused on the evolution of both communication and suppression of signaling in a competitive environment. They observed surprising evolutionary dynamics. While robots were quickly selected not to emit light on the food source (because it helped competitors find the food source), signaling was never completely suppressed. They found that it could be explained by the fact that the strength of selection for suppression decreased as the information in the signal diminished. Others have been interested in the evolution of swarming behaviours. As said earlier Olson & colleagues have studied the evolutionary mechanisms behind the emergence of herding behaviours (Olson et al. 2013a,b; Haley et al. 2014). While they categorize their work solely as IBM, it could be argued that their study belongs more precisely to the field of ER as they model evolution between embodied agents.

In consequence, evolutionary robotics allows to model individual variations and thus study their influence on the evolution of a given evolutionary trait. This implies that we can model the mapping from genotype to phenotype without making critical assumptions about the mechanistic constraints at play. Additionally, in comparison to IBM the embodiment that comes with ER decreases the assumptions we make on the exact nature of behaviours. This means that it may be possible for evolution to find unexpected solutions to coordinate which may not require any particular sensory or communication capabilities. This is appropriate as we do not want to limit the diversity of the possible behaviours which could evolve.

2.3.2 The Evolution of Cooperation in Evolutionary Robotics

As a final Section for this Chapter we review the works in ER that have been interested in the evolution of cooperation to present where we stand in this context. As is common when studying the evolution of cooperation, most works have been focused on altruism. Waibel & colleagues (Waibel et al. 2011) conducted an empirical test of Hamilton's rule³ (Hamilton 1964) for the evolution of altruism in a group of robots. They designed a foraging task where robots had the possibility to share with others their benefits obtained from foraging. They then tested how the coefficient of relatedness between robots influenced the evolution of altruism (i.e. sharing foraged resources). They showed that, in this context, Hamilton's rule is indeed quantitatively validated. Similarly in another study they have been interested in the influence of both the genetic composition of groups of robots (homogeneous or heterogeneous) and the level of selection (individual-level or group-level) in the evolution of cooperation (Waibel et al. 2009). In particular, they studied the impact of these two criteria on the performance of robots in three different foraging tasks: a solitary one, a cooperative one and an altruistic one. As could be expected, teams where individuals were homogeneous (i.e. where genetic relatedness

³Hamilton's rule is a way to summarize the effect of kin selection by stating that an altruistic trait may be selected if the following inequation is respected: $rb > c$, where b and c are respectively the fitness benefits on the recipient and costs on the actor of the cooperative interaction and r is the relatedness between the recipient and the actor. In consequence, an altruistic behaviour is favoured when the benefits of this behaviour weighted by the relatedness with the actor outweigh the cost of cooperating.

was equal to 1) performed better in tasks which required cooperation. Montanier and Bredeche have also studied the evolution of altruism in an *environment-driven* model. In this type of models, selection pressure comes solely from the environment. As such no fitness function is explicitly defined and individuals need to meet each other so that they can exchange genetic material (Bredeche and Montanier 2010), which is a more "realistic" approach to modeling evolution. In particular, they studied the evolution of altruism in a foraging setting under a "tragedy of commons" situation (Hardin 1968; Montanier and Bredeche 2011). This means that individuals have to share a common limited resource to such extent that some individuals may have to die so that the whole population does not go extinct. They showed that altruism can evolve under sufficient genetic relatedness. In another similar study (Montanier and Bredeche 2013), these authors validated the existence of a negative correlation between the evolution of altruism and spatial dispersion. More precisely, under low dispersion individuals tend to interact with other nearby individuals. Thus they interact with individuals that are more genetically related to them, which generates sufficient genetic relatedness so that kin selection can occur (Van Baalen and Rand 1998).

Others have been interested in the evolution of division of labour. These studies have in particular focused on the evolution of specialisation in ants. Ferrante et al. (Ferrante et al. 2015) proposed an ER model of task-partitioning (where a task has to be done in sequence) in leafcutter ants. In this species, some ants are tasked with cutting leaves and leave them in a storage location from which other ants collect the leaves and bring them back to the nest. They showed that division of labour could evolve when particular environmental features (in their case a slope) could be exploited to reduce switching costs. They thus validated a biological hypothesis about the role of switching costs in the evolution of specialisation (Duarte et al. 2011).

Solomon and colleagues (Solomon et al. 2012) are among the few who have been interested in modeling the evolution of cooperation between unrelated individuals, i.e. where cooperation is necessarily mutually beneficial (as altruism cannot evolve between unrelated individuals). They studied the evolution of signaling strategies in cooperative robots. In particular, they took inspiration from the hunting behaviours of spotted hyenas (Smith et al. 2012) in the context of competition against lions for the stealing of a prey. They compared the performance of two different signaling strategies: (1) one where all individuals can signal to the others and (2) another one where only a particular individual, the flag-bearer, may signal. The latter strategy was revealed to achieve higher coordination between individuals and therefore to increase the benefits of the cooperative action.

2.4 Conclusions

In conclusion of this Chapter, we do not claim any model to be fundamentally better than the others. As we previously stated in the general Introduction each model is based on assumptions and a specific level of abstraction. The choice of a model is made depending on which assumptions we expect to be of critical importance. Even considering a particular model to be globally more realistic can be tricky; some models simply represent

more accurately a particular aspect of a phenomenon. Mitri and colleagues (Mitri et al. 2013) classified models for the study of social behaviours according to their situatedness, which they defined as "the extent to which individuals are embedded in an environment that they can sense and modify". In this Chapter, our goal was to show that classical models in evolutionary biology make assumptions on the availability of mutations that we consider to be of importance in the evolution of mutualistic cooperation. In particular, we claim that the appearance of mutants (and thus the emergence of cooperation) depends on the proximate mechanisms of coordination. We thus motivated our choice of using individual-based modeling to address this issue. Thus we can model the mapping between genotype and phenotype and thus study the influence of mechanistic constraints on the evolution of cooperation. The framework of evolutionary robotics is relevant as we make no assumption on the nature of the behaviours evolved, which may allow for a higher diversity of coordination strategies.

In the next two Chapters, we will thus be interested in using ER to study the impact of the mechanics of coordination behaviours on the evolution of cooperation. First, we focus on comparing the evolution of collective hunting in a classical game theoretical model and in evolutionary robotics. We study the differences between these two approaches with regard to the emergence of a cooperative strategy. We thus investigate the role of coordination in the transition to cooperation. Then, in a second Chapter, we study more precisely how the nature of coordination strategies influence the evolution of collective actions. Namely, we are interested in the transition from a suboptimal collective equilibrium to the optimal one. We want to study how the evolution of different coordination strategies may impact the emergence of such transition thanks to individual selection only. In both of these studies, no physical robots are used. While it could allow to endow our studies with real physics (e.g. friction), we do not believe that level of realism to be of critical importance in this context. As with any model, we choose to abstract from some aspects of the real world that we think do not impact our results. Moreover, experiments on physical robots are in any case unfortunately too time-consuming to consider using them to such scope (Mitri et al. 2013; Doncieux et al. 2015).

Chapter 3

The Impact of Behavioural Mechanisms in the Evolution of Cooperation

Abstract	55
Introduction	55
Materials and Methods	57
Results	60
Discussion	64
Supporting Information	68

In this Chapter, we focus on the impact of proximate mechanisms in the evolution of mutualistic cooperation. This work is presented in the form of a published journal paper:

Arthur Bernard, Jean-Baptiste André, and Nicolas Bredeche (2016b). “To Cooperate or Not to Cooperate: Why Behavioural Mechanisms Matter”. In: *PLOS Computational Biology* 12.5, pp. 1–14

Our aim here is to show that the behavioural mechanisms of coordination are critical in the evolution of collective actions. In particular, we are interested in the evolution of mutually beneficial actions. As previously explained in the Introduction, mutually beneficial behaviours are stable once evolved because they benefit all of the individuals. However, their origin is not trivial because they often require the emergence of coordination. In consequence, no single individual may benefit from cooperation unless others also cooperate.

We take inspiration from the game theoretical model of the stag hunt (Skyrms 2004). We claim that crucial assumptions are made in classical theoretical models which may hide the complexity of evolving mutualistic actions. In particular, it is often assumed that a single mutation is sufficient to evolve a cooperative individual from a solitary individual. In reality, the evolution of cooperation entails the emergence of multiple traits that are not

beneficial on their own. We thus believe that the impact of mechanistic constraints on the evolution of coordination has been neglected by these classical models. In order to take these mechanisms into account, we model the stag hunt in evolutionary robotics (Nolfi and Floreano 2000; Doncieux et al. 2015). Two individuals evolve in an enclosed arena filled with 18 prey, half of them being hares, the other half being stags. Hunting a hare rewards less than hunting a stag and can be achieved in a solitary or cooperative fashion. In comparison, the benefits of hunting a stag cannot be reaped unless both individuals capture the prey together.

In this Chapter, we reveal drastic differences when the transition to stag hunting is modeled in evolutionary robotics. In particular, we show that the evolution of the cooperative equilibrium (i.e. stag hunting) when starting from initially solitary individuals (i.e. hare hunting) always occurs with a classical game theoretical model. In comparison, in evolutionary robotics the evolution of cooperation is nearly impossible as it happens in only 1 replication out of 30. Moreover, even when individuals are genetically related, the transition to cooperation is still unlikely as it occurs in 20% of all replications. Our model thus reveals that the transition to cooperation is faced with a chicken and egg dilemma. More precisely, cooperation cannot be selected unless it is beneficial for the individual. Yet the benefits of cooperation cannot be reaped unless other individuals are capable of coordination. In particular, we observe that the evolution of coordination requires the emergence of a complex behaviour which is unlikely to evolve on its own. In consequence, we demonstrate that the practical mechanics of coordination deeply impact the evolution of mutually beneficial cooperation. We thus argue that models that consider the proximate explanations of cooperation are crucial in order to grasp a full understanding of the evolution of mutualistic cooperation.

To Cooperate or not to Cooperate: why Behavioural Mechanisms Matter

Abstract

Mutualistic cooperation often requires multiple individuals to behave in a coordinated fashion. Hence, while the evolutionary stability of mutualistic cooperation poses no particular theoretical difficulty, its evolutionary emergence faces a chicken and egg problem: an individual cannot benefit from cooperating unless other individuals already do so. Here, we use evolutionary robotic simulations to study the consequences of this problem for the evolution of cooperation. In contrast with standard game-theoretic results, we find that the transition from solitary to cooperative strategies is very unlikely, whether interacting individuals are genetically related (cooperation evolves in 20% of all simulations) or unrelated (only 3% of all simulations). We also observe that successful cooperation between individuals requires the evolution of a specific and rather complex behaviour. This behavioural complexity creates a large fitness valley between solitary and cooperative strategies, making the evolutionary transition difficult. These results reveal the need for research on biological mechanisms which may facilitate this transition.

Introduction

It is well known that, in the absence of genetic relatedness, altruistic behaviours in which individuals pay a fitness cost for the benefit of others cannot evolve by natural selection (Hamilton 1964; West et al. 2007b). However, it is often assumed that mutualistic behaviours, wherein individuals collectively gain a common benefit (Leimar and Connor 2003; Leimar and Hammerstein 2010), do not pose such a problem, and are therefore of limited interest to evolutionists: they simply evolve because they benefit the individuals who express them.

However, mutualistic behaviours do often pose a different kind of evolutionary problem than altruism: they require coordination (Alvard and Nolin 2002; Alvard 2003; Drea and Carter 2009; Leimar and Connor 2003). Many collective traits are only mutually beneficial if several individuals express them together in a coordinated fashion. That is, it would not be beneficial for a single individual to express the cooperative trait if others did not express it as well. Consequently, whereas altruistic behaviours pose a problem of *stability*, which can only be solved by genetic relatedness, many forms of mutualistic behaviours pose a problem of *evolution*. These collective strategies are stable equilibria but their evolution is complex.

This problem has been formalized in game theory as the stag hunt game (Skyrms 2004). In the stag hunt, two hunters are confronted with the choice of either hunting a hare alone for a small but guaranteed benefit, or coordinating to hunt a stag cooperatively for a bigger

reward, with the risk of not being rewarded at all if they hunt the stag alone. There are two evolutionarily stable Nash equilibria in this game: (1) simultaneous defection (i.e. both players hunt hares), which is risk-dominant as it maximizes the minimum payoff an individual can expect, and (2) simultaneous cooperation (i.e. both players hunt stags), which is payoff-dominant as it maximizes the total payoff at equilibrium. One of the aims of evolutionary analyses of the stag hunt is to characterize the mechanisms that facilitate the transition from the solitary equilibrium to the cooperative equilibrium. The difficulty is that cooperation can only be favoured by selection when a sufficient proportion of individuals in the population also cooperate. The transition from a population with a majority of solitary individuals to one with a majority of social individuals requires the rise of cooperation above an invasion threshold, which must occur for non-selective reasons.

In game-theoretic analyses, the hunting strategy of individuals is generally assumed to be encoded by a single genetic locus with two alleles: solitary or social (Skyrms 2004). In this case, random mutations and/or demographic stochasticity can lead to the appearance of a subpopulation of mutants playing the social strategy which is sufficient to overcome the invasion threshold. Moreover, Skyrms (Skyrms 2004) showed that this cooperation can be further facilitated in a spatially structured population in which individuals tend to interact more with genetically related partners.

However, this approach makes a very strong assumption about the underlying mechanistic nature of behaviour: that a single mutation is sufficient to transform an individual playing a solitary strategy into an individual playing a perfectly efficient social strategy. In reality, hunting socially implies several novel behavioural abilities. In particular, it implies the ability to coordinate with others in order to focus on the same prey, which is unlikely to occur with only a single random mutation. In this Chapter, we postulate that critical aspects of coordinated cooperation have been neglected by game-theoretic analyses and investigate the mechanistic constraints which interfere with the evolution of coordination in a more realistic setting where the mapping between genotype and phenotype is not limited to a strict binary encoding.

Evolutionary robotics is a useful methodology for the simulation and study of this more realistic conception of behaviour and its genetic underpinnings (Nolfi and Floreano 2000; Doncieux et al. 2015). This approach allows to simulate the evolution of complex genotypes and observe the resulting behaviours in robotic agents. Such simulations also make it possible to investigate the complex mechanistic constraints at play in the translation from genotype to phenotype (Mitri et al. 2013). A considerable body of work has already been dedicated to modeling social evolution with robotic approaches (Trianni 2014). These studies have been interested in a large diversity of issues: the evolution of swarms (Olson et al. 2013a), the mechanics of division of labour in social insects (Tarpore et al. 2010; Ferrante et al. 2015) or the evolution of communication (Floreano et al. 2007; Mitri et al. 2011; Wischmann et al. 2012; Solomon et al. 2012). The evolution of cooperation in particular has been addressed in numerous papers. In the vast majority of this literature, however, social partners are genetically related (Waibel et al. 2009), whether motivated by design (Hauert et al. 2014; Trianni et al. 2007) or to study the evolution of altruism (Waibel et al. 2011; Montanier and Bredeche 2013). Few articles,

in comparison, have been interested in the evolution of mutualistic cooperation between genetically unrelated individuals (Solomon et al. 2012). Moreover the specific problem posed by the stag hunt game, where cooperation is not the only evolutionarily stable strategy and a non-collective solution acts as a stable attractor, has never been studied in evolutionary robotics.

In this Chapter, we use an experimental model where simulated robotic agents interact in a situation equivalent to the stag hunt and compare the results of our model to those of standard game-theoretic analyses. Our results shed new light on the influence of mechanistic constraints in the evolution of coordinated actions. We then use this model to explore realistic mechanisms that could drive the transition to collective behaviours.

Materials and Methods

Experimental Setup

We consider an environment with two hunters and several prey, both hares and stags. Hunters can choose to hunt either of these prey, earning different food rewards depending on whether they hunt alone or cooperate (see Fig. 3.1).

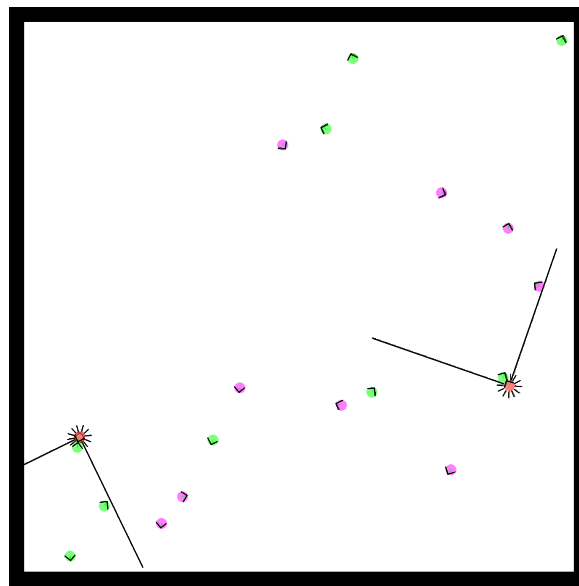


Figure 3.1 – *Screenshot of a robotic simulation.* The red dots represent the two hunters, the green dots the hares, and the pink dots the stags. The black lines around the agents' body represent the proximity sensors and the black cones on front the cameras described in the text. Hunters are allowed to move throughout the environment. Hares and stags remain at their starting positions.

Food rewards for killing a prey are shown in Table 3.1. A hare yields a reward of 50, regardless of whether it is hunted in a solitary or cooperative fashion. A stag yields a reward of 500 for each hunter only if it is hunted cooperatively. If a stag is killed by a

single hunter, it is still removed from the arena but is considered a failed hunt and rewards nothing. None of the rewards are split between cooperators.

Prey		Food Reward
Hare	<i>alone</i>	50
	<i>coop.</i>	50
Stag	<i>alone</i>	0
	<i>coop.</i>	500

Table 3.1 – *Food rewards for hunting different prey.* The reward depends on whether these prey were hunted alone or cooperatively. There is no reward for stags hunted alone in this case.

Simulated robotic agents are evaluated in an 800 by 800 unit square arena, which has four solid walls and is devoid of any obstacles aside from other agents. Each circular-shaped agent, with a diameter of 14 units, is equipped with two independent wheels and a collection of sensors. Hunters can use the information provided by 12 proximity sensors and a front camera. Proximity sensors have a range of approximately twice the diameter of the agent’s body, and provide the agent with the proximity of the nearest obstacle. They are evenly distributed around the agent’s body. The front camera consists of 12 rays with infinite range spread out in a 90 degree cone in front of the body. Each ray in the camera provides two different pieces of information about the first target it intersects with: the type of target (hunter, hare, or stag) and its proximity. This robot model facilitates the evolution of basic walls avoidance and agents recognition behaviours, which we consider not to be of interest here. Hence we separate obstacles recognition (by the proximity sensors) from agents’ recognition (by the camera).

Only the hunters are capable of movement; prey remain at their initial positions. (Complementary experiments with moving prey capable of avoidance behaviours did not produce significantly different results; not shown.) A prey is caught if any hunter remains close enough during a fixed amount of time steps (800 steps, in a simulation lasting 20.000 time steps). Cooperative hunting is defined as a prey with two hunters in catching distance at the time of its capture. Therefore, cooperation happens even if only one of the two hunters is in catching distance of the prey for most of the time, as long as the two hunters are there in the final step. The prey is then immediately replaced at a random position in the arena, thus keeping a fixed number of agents and prey during the whole simulation.

Neural Network for Agent Control

The hunters’ behaviour is computed by an artificial neural network which maps sensory inputs to motor outputs. The neural network is a fully connected multi-layer perceptron with a single hidden layer of 8 neurons. The inputs of this network are the perceptions of the agent, with 12 neurons for the proximity sensors and 48 for the camera (4 for each of the 12 rays) plus a bias neuron (whose value is always 1), for a total of 61 input

neurons. The two outputs of the network control the speed of each of the agent's wheels and the mapping function between inputs and outputs is a sigmoid function (see Fig. 3.2). Changing the number of hidden neurons did not yield significantly different results (not shown).

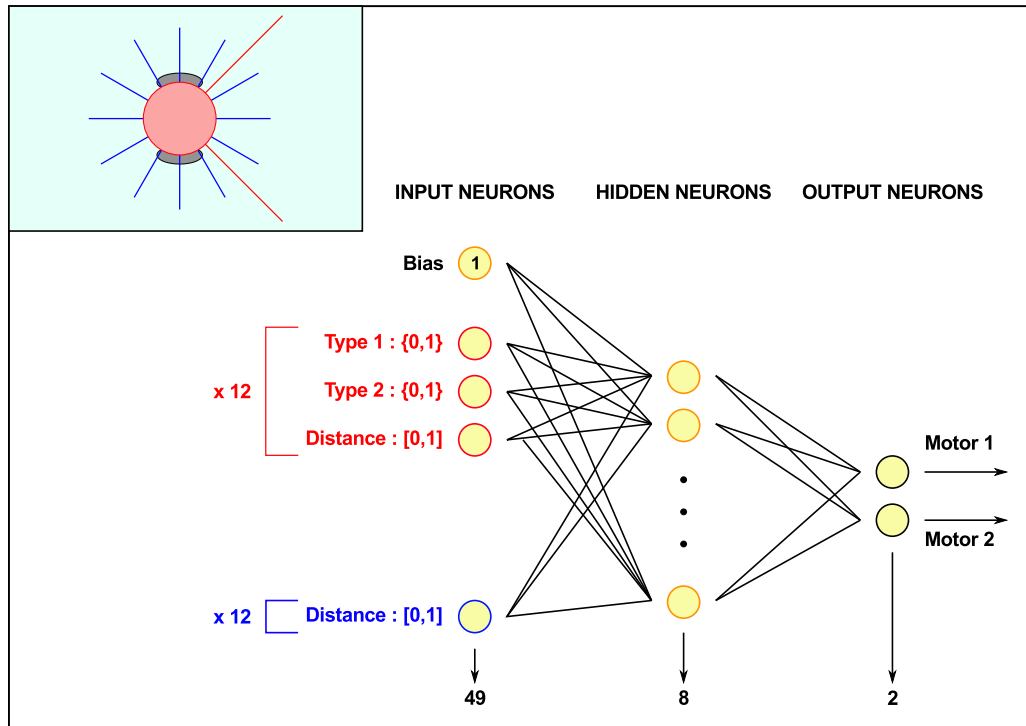


Figure 3.2 – *Diagram of the simulated robotic agent used in the simulation (inset) and its neural network controller.* The blue lines represent the 12 proximity sensors and the red lines represent the front camera. Inputs "Type 1" and "Type 2" are two boolean values used to represent the type of the agent (encoded with two bits) recognized by the camera ray.

Simulating Artificial Evolution

To simulate evolution, we use an evolutionary algorithm to evolve the genome of the hunters. This genome is comprised of a collection of 410 real values in the range $[0, 1]$, one for each of the neural network's weights, and is initially randomized for each individual in the population. In order to obtain its fitness, each individual is successively paired five times with a partner randomly chosen each time (except itself) in the arena presented in the Experimental Setup section, for an evaluation round of 20.000 time steps. The payoff of the evaluated individual at the end of a round is given by the total amount of food it has managed to obtain by killing prey in this round. As this quantity depends heavily on the initial conditions (random initial positions of the prey), five simulations are performed for each pair of individuals. The individual's fitness is then obtained by computing the sum of payoffs averaged over the total number of simulations for the individual. In this case the number of simulations is 25, with 5 partners and 5 simulations with each partner.

Experiments were conducted using a Wright-Fisher model (Wright 1931) with constant population size (20 individuals), which is commonly known as a fitness-proportionate selection method in evolutionary robotics (Eiben and Smith 2015). Using this model, the population of the next generation is formed by a random sampling of offspring from the previous generation, with the probability of sampling a particular parent proportional to the parent’s fitness. Each offspring is simply a mutated clone of its parent; recombination is not included in our simulation. Consequently, new genotypes appear only through mutation. These mutations are performed using a Gaussian function, with a standard deviation of 2×10^{-1} and a mutation probability of 5×10^{-3} . Each experiment lasted 3000 generations. All simulation parameters are summarised in Table 3.2.

Parameter		Value
Evolutionary Algorithm		
	Selection method	Fitness-proportionate
	Population size	20
	Gene mutation probability	5×10^{-3}
	Mutation operator	Gaussian $\mathcal{N}(0, 0.01)$
	Number of partners	5
	Number of simulations per pair	5
Artificial Neural Network		
	Input neurons	61
	Hidden neurons	8
	Output neurons	2

Table 3.2 – *Simulation parameters.*

Results

Starting with a population of hare hunters

In order to explore the evolutionary transition between the risk-dominant equilibrium (hare hunting) and the payoff-dominant equilibrium (cooperative stag hunting), individuals first evolved in an environment composed solely of hares. This ensured that the populations initially reached the solitary equilibrium. Only then did we add stags and study the dynamics of evolution. Fig. 3.3(A) shows the evolution of the mean percentage of stags hunted successfully (i.e., hunted cooperatively) out of the total number of prey hunted over time for 30 independent runs. Fig. 3.3(B) shows the mean proportion of each type of prey hunted during the last generation of each run. Stag hunting evolved in only one run out of 30 and even in that run accounted for less than 30% of the total number of prey hunted. In the other 29 runs, the individuals hunted only hares as they had previously evolved to do. These simulations demonstrate that the evolution of collective hunting is very unlikely when the population is composed of individuals who are already efficient solitary hunters.

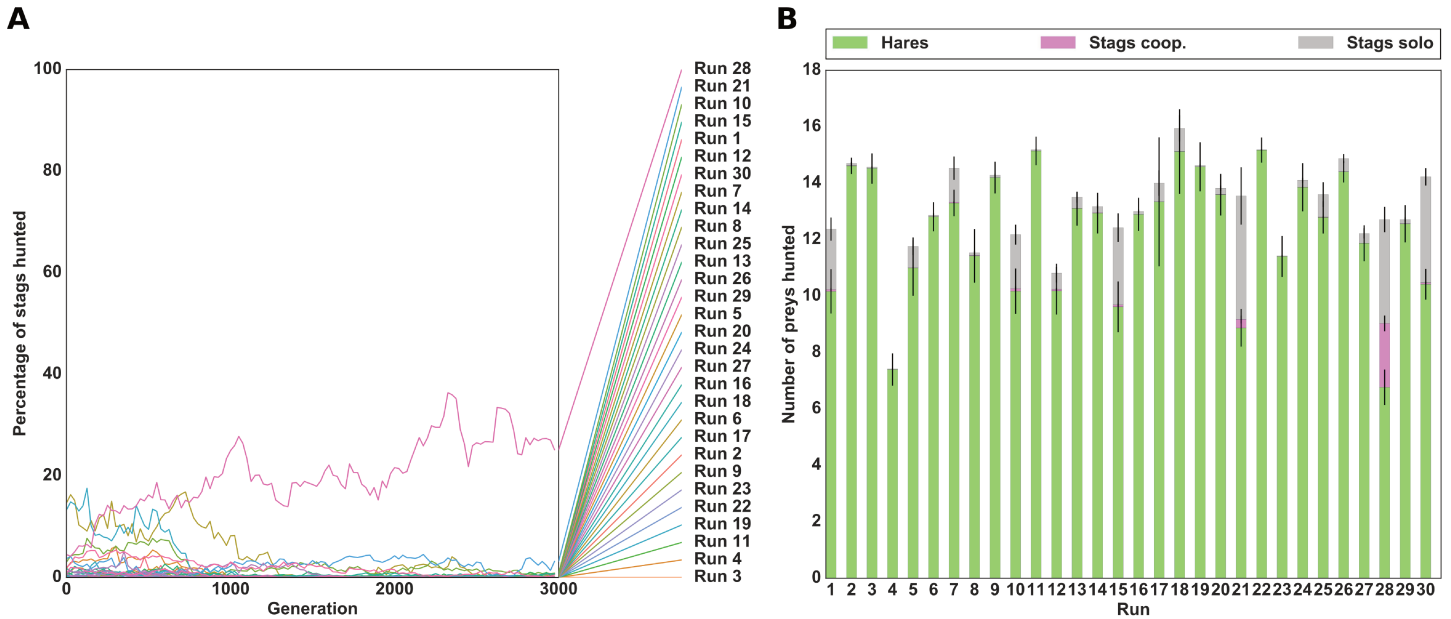


Figure 3.3 – Evolution of cooperation in a robotic simulation with an initial hare-hunting strategy. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted during the last generation of evolution for each independent run. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted successfully (cooperatively), and the top grey bar the number of failed hunts (stags hunted alone). The standard deviation for each quantity is shown by black lines. The population for each of the 30 independent runs was previously evolved in an environment with only hares. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.

For comparison we simulated the same scenario using the standard game-theoretic version of the stag hunt, where the expression of the two types of behaviour was encoded by a single binary locus. Each individual in the population initially possessed the allele for hare hunting (Fig. 3.4).

Here the transition to collective hunting occurred in each of the 30 independent runs and this strategy then remained stable. This result differs drastically from the results of our robotic simulations in which this transition never fully occurred (Mann-Whitney U test on the proportion of stags hunted successfully during the last generation, p -value < 0.001).

Starting with a random initial population

In a second experiment, we wanted to investigate the evolution of hunting strategies "from scratch", with the individuals' genotypes initialized with random values, rather than evolved with a specific hunting strategy. Fig. 3.5 shows the mean percentage of

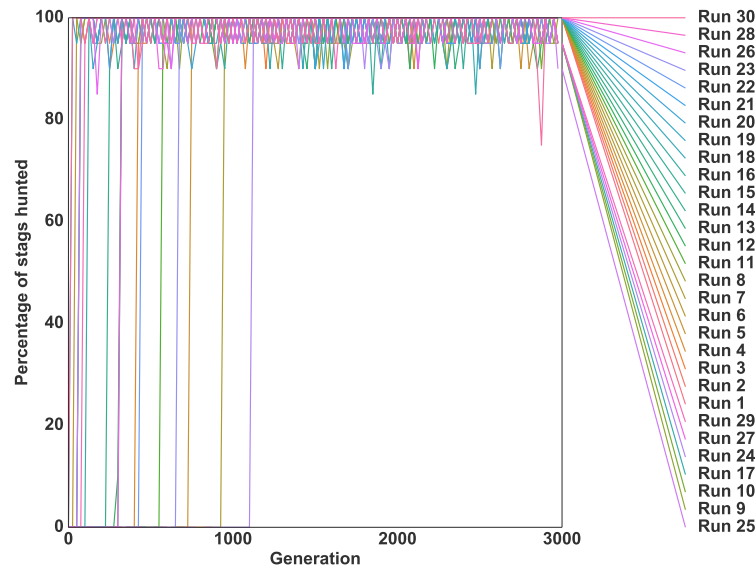


Figure 3.4 – Evolution of cooperation in a game-theoretic simulation with an initial hare-hunting strategy. Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted when starting with a population of hare hunters for 30 independent runs. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1.

stags hunted over time and the mean number of prey hunted during the last generation. We observed the transition to a clearly cooperative strategy in a single run, while in two other runs, 50% of prey hunted were stags. In the 27 remaining runs the proportion of stags hunted was less than 25%. In comparison, in simulations using the standard game-theoretic version of the stag hunt where individuals are initially unable to hunt, stag hunting evolved and remained stable in every run (see supporting information, Fig. 3.9).

The above experiments show that mechanistic constraints have a critical effect on the evolution of coordinated collective actions. In a simple game-theoretic analysis in which the hunting strategy is encoded by a single binary gene, collective behaviour systematically evolved. However, in a setting where the hunting strategy was determined by a more complex artificial neural network, cooperative behaviour evolved in fewer than 10% of cases. These results encourage further exploration into the evolutionary origin of coordinated collective actions and the mechanisms which may facilitate their evolution. In the following section, we explore two such mechanisms.

When stags can be hunted alone

In the next experiment, food was also rewarded for hunting a stag in a solitary fashion so that cooperative behaviour did not entail a risk. We wanted to study whether hunting a stag alone could act as a transition towards the evolution of the collective strategy. Hunting a stag alone was given the same reward as hunting a hare (Table 3.3), differing from classical models of the stag hunt.

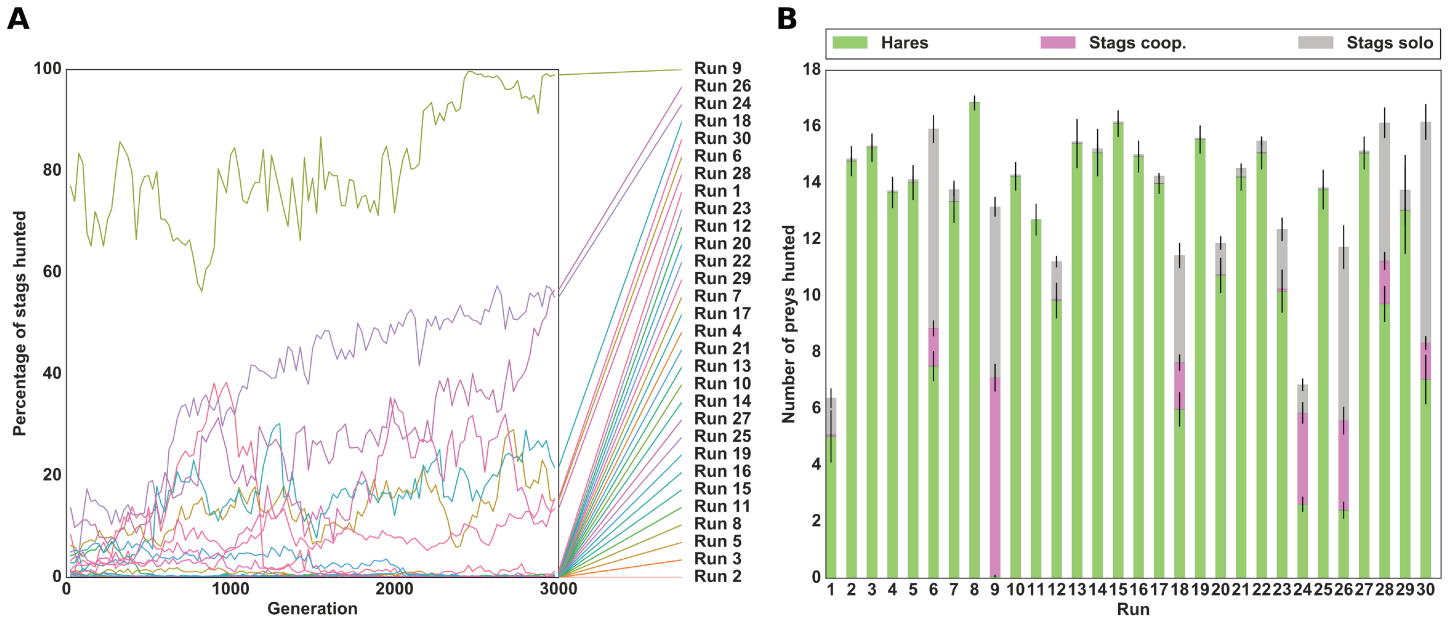


Figure 3.5 – Evolution of cooperation with no initial hunting strategy. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted successfully (cooperatively) and the top grey bar the number of failed hunts (stags hunted alone). The standard deviation for each quantity is shown by black lines. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively, as presented in Table 3.1. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.

Fig. 3.6 shows the results of robotic simulations where individuals initially evolved to hunt hares (as in Fig. 3.3). As expected, the evolution of collective hunting was significantly facilitated when the risk of hunting stags alone was removed (Mann-Whitney, p -value < 0.001). The populations completely switched to hunting stags in two runs out of 30, and in three other runs, more than 50% of the prey hunted were stags, with a large part of the prey hunted cooperatively in each of these runs. However, in most of the runs (25 out of 30), the evolved strategy was to hunt both types of prey in a solitary fashion. From these results, it entails that the individuals are still hindered by the evolution of a successful coordination strategy.

The role of genetic relatedness

Genetic relatedness among social partners is known to influence the evolution of many types of social traits (Hamilton 1964). In particular, (Skyrms 2004) showed how it can facilitate the evolution of cooperation in a stag hunt game (Skyrms 2004, chapter 3). It can yield more frequent interaction between cooperators, which in turn increases their probability of benefiting from cooperative behaviour. In order to include this mechanism,

Prey		Food Reward
Hare	<i>alone</i>	50
	<i>coop.</i>	50
Stag	<i>alone</i>	50
	<i>coop.</i>	500

Table 3.3 – Food rewards for hunting different prey. *The reward depends on whether these prey were hunted alone or cooperatively. There is a reward for stags hunted alone in this case.*

we considered an extreme situation in which each individual is always paired with a clone of itself, known as "clonal selection" in robotics, ensuring a maximal genetic relatedness of 1.

These results show that genetic relatedness has a positive effect on the evolution of cooperation (Fig. 3.7). In four out of 30 runs the population evolved the cooperative strategy. Moreover, in two other runs, stags accounted for more than 75% of prey hunted, as compared to less than 25% without relatedness (Mann-Whitney, p -value < 0.005). When the initial population was random, rather than only hare hunters (see supporting information, Fig. 3.10), the positive effect of genetic relatedness was also observed in 12 out of 30 runs, where more than 50% of prey hunted were stags.

Discussion

There is a profound difference between evolutionary game-theoretic and robotic simulations of the stag hunt. Using identical model parameters, the transition from the solitary equilibrium to the social equilibrium always occurred in game-theoretic simulations, but was extremely unlikely in robotic simulations, occurring in 1 run out of 30. The complexity of the mapping between genotype and phenotype is responsible for much of this contrast. Individuals involved in a coordination game such as the stag hunt face a chicken & egg problem: the cooperative behaviour must be beneficial in order to evolve, but no individual can benefit from this behaviour unless the behaviour is already expressed by other individuals. When binary variation at a single genetic locus encodes the expression of the solitary or cooperative strategy, a single mutation is sufficient for a cooperative mutant to appear in a resident population of solitary individuals. In a finite population, demographic stochasticity can then lead to the rise of cooperators above the invasion threshold, at which point natural selection leads to their fixation, switching from a solitary equilibrium to a social one. In contrast, in our robotic simulations, the mapping between genotype and phenotype is more complex. Adopting the social strategy entails both a modification of the preferred hunting target and the ability to coordinate with others. Thus, several mutations are necessary for the appearance of full-fledged cooperative behaviour. As several individuals must carry these multiple mutations for the behaviour to become beneficial, the transition to the cooperative equilibrium is nearly impossible.

In particular, in our robotic simulations we were able to observe that coordination

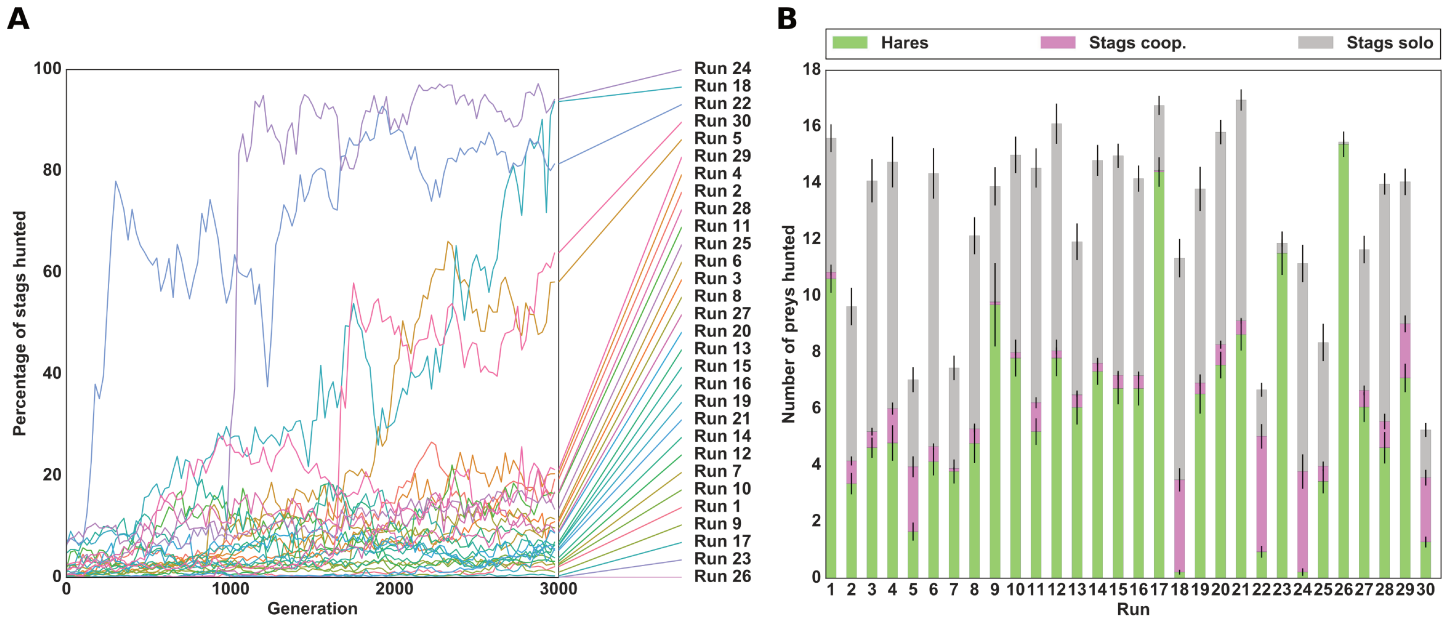


Figure 3.6 – Evolution of cooperation with an initial hare-hunting strategy and a reward for solitary stag hunting. (A) Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted cooperatively and the top grey bar the number of stags hunted alone. The standard deviation for each quantity is shown by black lines. The population for each of the 30 independent runs was previously evolved in an environment with only hares. Rewards were 50 for a hare, 50 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.3. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.

entails a specific and rather complex behaviour. Fig. 3.8 shows the behaviours evolved by the best individuals in the cooperative run shown in Fig. 3.5 (Run 9). The solution they evolved for coordination was to circle around one another, allowing each of them to constantly see their partner while both moving closer to a stag. This behaviour was replicated in every cooperative run. We thus observed the evolution of an ingenious (given the agents' limited capabilities) and complex hunting strategy. These findings demonstrate that the practical mechanics of behaviour can have important evolutionary consequences, and that models which ignore these properties may lead to misleading predictions.

Moreover, the evolution of cooperation is also strongly impacted by ecological features. Social hunting poses a bootstrapping problem because it entails both a modification of the preferred hunting target and an ability to coordinate with others. Its evolution can be facilitated, therefore, if hunters have a reasonable probability of hunting the same prey as their partner, just by chance, with no need of active coordination. Biologically, this could occur if hunters live in a dense social environment (with many other hunters in the

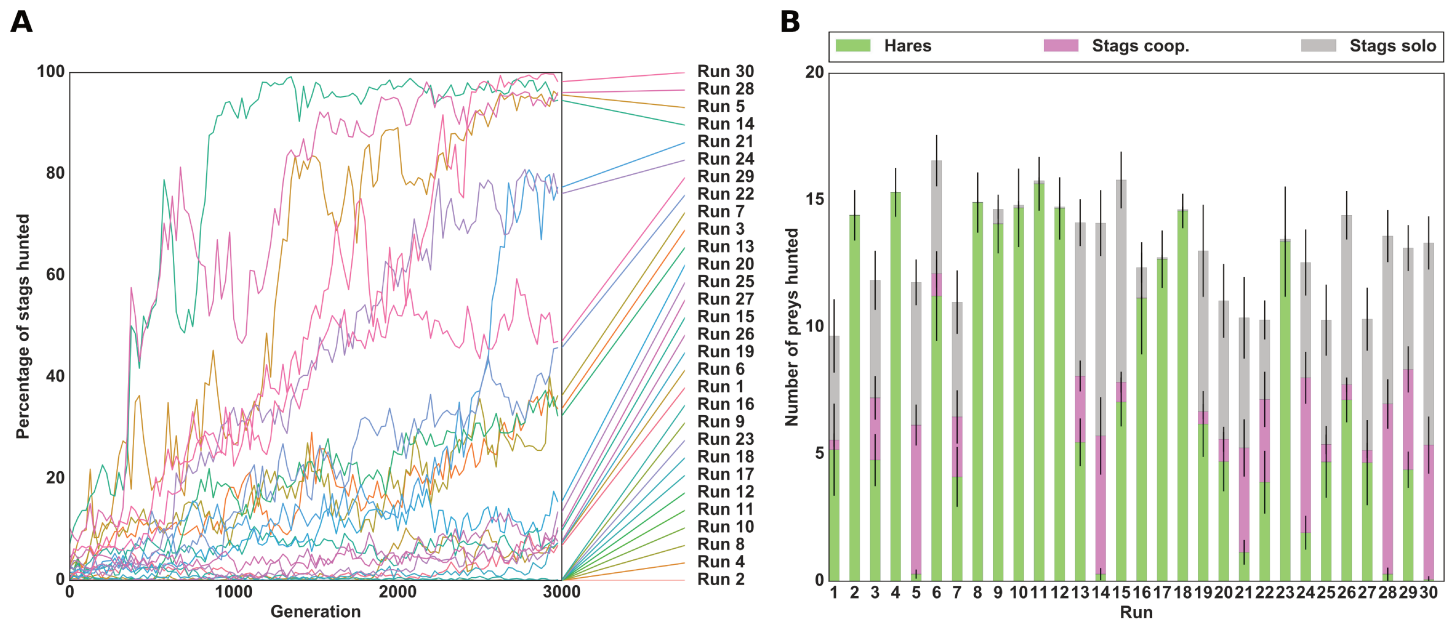


Figure 3.7 – Evolution of cooperation under maximal genetic relatedness with an initial hare-hunting strategy. (A) Evolution of the mean percentage of stags hunted successfully (*i.e.* cooperatively) with respect to the total number of prey hunted in a robotic simulation. (B) Mean number of prey hunted during the last generation of evolution for each independent run. The bottom green bar represents the number of hares hunted, the middle pink bar the number of stags hunted successfully (cooperatively) and the top grey bar the number of failed hunts (stags hunted alone). The standard deviation for each quantity is shown by black lines. The population for each of the 30 independent runs was previously evolved in an environment with only hares. The genetic relatedness between paired individuals was 1. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1. The number of prey (18) was kept constant throughout the simulation by replacing killed prey by a prey of the same type.

vicinity), and/or if the density of prey is low, such that the likelihood of ending up on the same prey is large. To test this possibility, we conducted additional experiments where the density of prey was varied. The number of prey was whether (1) decreased from 18 to 6 or (2) increased from 18 to 30. The population was initially constituted of hare hunters and we kept the same ratio of prey as in previous experiments (*i.e.* 50% of hares and 50% of stags). We show (see supporting information, Fig. 3.11(A)) that when the number of prey is decreased (6) the transition to a cooperative strategy is facilitated (Mann-Whitney, p -value < 0.05) as in 9 runs out of 30, more than 30% of the prey hunted are stags. In comparison, a higher density of prey (30) entails that it is impossible to evolve cooperation (see supporting information, Fig. 3.11(B)). These results reinforce our claim that the practical mechanics of coordination are crucial in understanding the evolution of cooperation. In particular, here, the precise ecological situation faced by individuals plays a key role in the transition to the collective equilibrium.

Finally, the complexity of coordination suggests that the recycling of a previously

evolved trait could be necessary for the transition to cooperation, i.e. individuals could coordinate thanks to behavioural features that may not have been selected for cooperation at first. Such features could include the evolution of communication, or a leader-follower strategy. The role of both of these behaviours has already been studied in real-life stag hunt type interactions in chimpanzees and human children (Bullinger et al. 2011; Duguid et al. 2014), and there is an already extensive literature in evolutionary robotics on their role in the evolution of collective actions (Trianni et al. 2007; Mitri et al. 2009; Solomon et al. 2012; Ferrante et al. 2015). This offers some directions for future works on this problem.

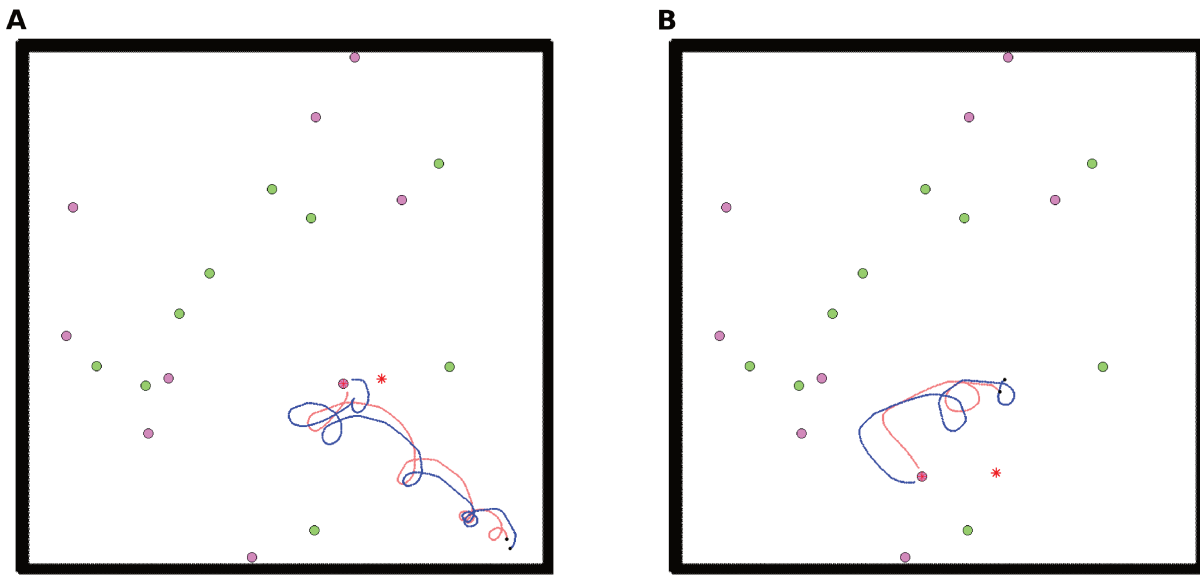


Figure 3.8 – Snapshots of a simulation after two hunts. In each of these snapshots, we show the path travelled by each hunter (in different colours) since their last prey was hunted. The black dots represent the positions of the hunters at their last kill. The red star on the stag (pink circle) converged on by the hunters indicates not only that the prey was killed but, more importantly, that it was killed cooperatively by the two hunters.

Supporting Information

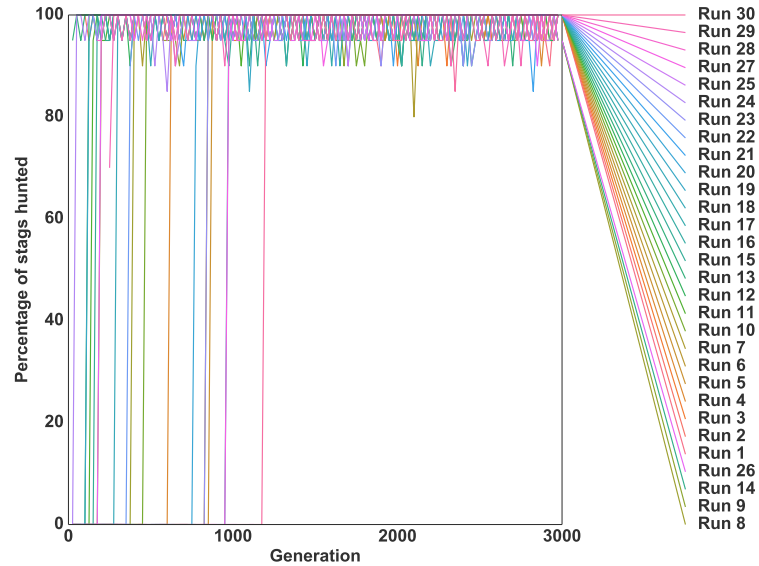


Figure 3.9 – Evolution of cooperation in a game-theoretic simulation with an initial hare-hunting strategy. Evolution of the mean percentage of stags hunted with respect to the total number of prey hunted where individuals are initially unable to hunt for 30 independent runs. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively.



Figure 3.10 – *Evolution of cooperation under maximal genetic relatedness with no initial hunting strategy.* Evolution of the mean percentage of stags hunted with respect to the total number of prey hunted in a robotic simulation. The genetic relatedness between paired individuals was 1. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented.

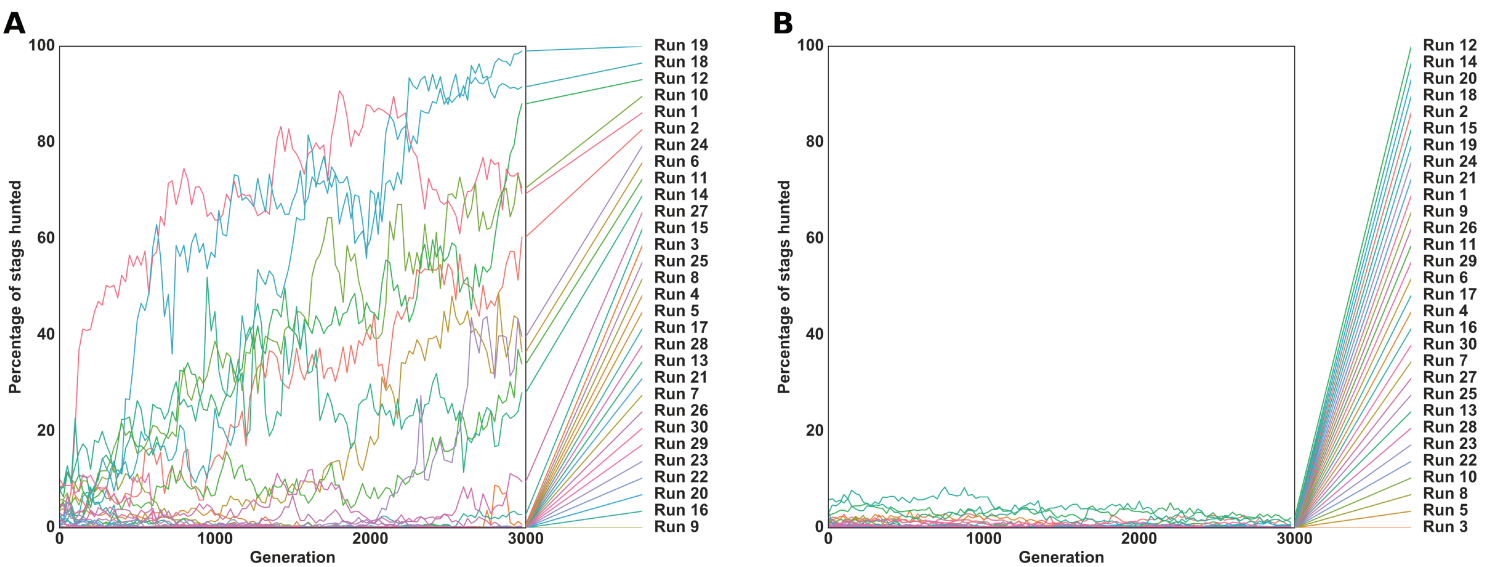


Figure 3.11 – *Evolution of cooperation with an initial hare-hunting strategy and a varied density of prey.* Evolution of the mean percentage of stags hunted successfully (i.e. cooperatively) with respect to the total number of prey hunted in a robotic simulation when the number of prey was (A) 6 and (B) 30. The population for each of the 30 independent runs was previously evolved in an environment with only hares. Rewards were 50 for a hare, 0 for a stag hunted alone, and 500 for a stag hunted cooperatively as presented in Table 3.1. The number of prey was kept constant throughout the simulation by replacing killed prey by a prey of the same type.

Chapter 4

The optimization of Collective Actions by Individual Selection

Abstract	73
Introduction	73
Materials and Methods	75
Results	76
Discussion	82

In this Chapter, we investigate how the nature of coordination behaviours influences the optimization of collective actions. This Chapter is presented as a draft for a journal article.

In the previous Chapter we revealed that the evolution of collective actions was hindered by the evolution of coordination. In particular, we used evolutionary robotics to shed a light on the mechanistic constraints at play in the transition from a solitary to a cooperative equilibrium. Here we are interested in understanding how individual selection can lead to the transition between different collective equilibria.

Collective actions reap their benefits through the interactions between multiple individuals. However while they may benefit every individual in a mutualistic fashion, it is not clear how these collective behaviours are reached. More precisely, because they require the coordination of several individuals, multiple stable equilibria can emerge. Additionally, because benefits are reached through a collective action, a single individual deviating from the evolved equilibrium would not be favored by selection. In other words, a mutant acting toward a different collective equilibrium would not be selected, even if the equilibrium is more advantageous for the group. In consequence, the issue of the optimization of collective actions arises. Namely we wonder how it is possible for individual selection to lead the transition toward an optimal equilibrium when another collective equilibrium already evolved ?

One classical mechanism to solve this issue is group selection. Because those behaviours are beneficial at the level of the group then selection could only occur at the same level,

so the argument goes. Here we hypothesize that collective behaviours can be optimized by individual selection only. To that end, we model the example of collective hunting. Individuals evolve in an environment where they can hunt two differently rewarding types of prey: *boar* and *stag*. In comparison to the hare of stag hunt, the boar can be hunted in a solitary fashion but rewards more when hunted cooperatively. Each type of prey corresponds to a different collective equilibrium: suboptimal for the boar and optimal for the stag. Our goal is thus to study the transition from the suboptimal equilibrium (i.e. boar hunting) to the optimal equilibrium (i.e. stag hunting).

We reveal that under simple ecological features where only two prey are present in the environment, the transition to the optimum is impossible. However, under more realistic assumptions where the individuals have to choose between multiple prey, then the optimal equilibrium evolved in 8 replications out of 30. In particular, the individuals now have to coordinate in order to achieve cooperation. This means that they need to react to each other's behaviour. From this it stems that they also react to a mutant's behaviour. This in turn may allow the group to reap the benefits of stag hunting.

However, in the collective strategy evolved by the individuals, they both separately decide which prey to hunt, leading to weak coordination. We then study how a more asymmetrical coordination strategy could impact the transition to the optimum. To that end, we increase the complexity of the artificial neural networks controlling the individuals in order to allow them to evolve more complex coordination strategies. In this case, we reveal that the transition to the optimum is facilitated as stag hunting evolves in 24 replications out of 30. Furthermore, we observe the evolution of a more efficient asymmetrical strategy where the individuals adopt two different roles: the *leader/follower* strategy. In this strategy, only the leader decides on which prey to hunt and the follower goes on the same prey. In consequence, while choosing to cooperatively hunt a stag was previously a collective decision making problem, it is now an individual problem. This means that a mutant leader going for a stag is now sufficient for both individuals to reap the benefits of stag hunting. Moreover, the leader/follower strategy evolved because it was more efficient for each individual. Thus we show that the evolution of an individually adaptive coordination strategy may lead to the optimization of a collective behaviour.

The Optimization of Collective Actions by Individual Selection

Abstract

Many social behaviours lead to the coexistence of multiple stable equilibria. However, a single individual deviating from a given collective equilibrium cannot be favored by selection. In consequence the transition to the optimal equilibrium when another collective equilibrium already emerged is commonly explained by group selection. Here we study the optimization of group-traits thanks to individual selection. We focus on a problem of collective hunting, where individuals have to choose between a suboptimal and an optimal prey, which we model in evolutionary robotics. We reveal that while the switch to the optimum is impossible under simple environmental conditions, it can occur when the environment is more complex. In particular, when coordination is necessary to cooperate, individuals react to each other's behaviour. As such a mutant's behaviour can affect that of the group and lead to the optimum. We then reveal that when a more efficient coordination strategy is evolved, the transition to the optimum is facilitated. Furthermore, the evolution of this new strategy is individually beneficial and is not enforced on the individuals. In consequence, we show that the optimization of collective actions can occur by individual selection.

Introduction

Many social traits are beneficial thanks to the emergence of collective features. For instance, collective hunting may allow to kill stronger or faster prey than what could be done in a solitary fashion. However this behaviour is only individually adaptive in interaction with a group of other collective hunters. The benefits of these behaviours are thus reached through the emergence of a group-trait. This raises the issue of how they could evolve under individual selection.

Once a collective strategy has emerged, a single mutant following a different strategy would not be favoured by selection. As such multiple evolutionarily stable equilibria may exist, as is the case in coordination games (Maynard Smith 1982; Skyrms 2004). Yet it is not clear how to predict which equilibrium will emerge. In particular, individuals may evolve a suboptimal equilibrium instead of the optimal one. Therefore, because no individual can benefit by deviating from the equilibrium evolved by the group, this raises the issue of optimizing the collective behaviour. Namely, how can the optimal equilibrium be reached when a suboptimal equilibrium has already emerged ?

A popular mechanism to explain the selection of equilibria is that of group selection (Boyd and Richerson 1990). Because these collective traits are adaptive at the level of the group it seems that only a selective pressure acting at the same level can lead to their optimization. Different groups will reach various collective equilibria and thus benefit differently from the evolved equilibria. In consequence, the groups for which a higher

payoff equilibrium emerged will be at a selective advantage in comparison to others and their strategy is thus expected to spread in the population. The evolution of cultural norm in particular is a popular group-level explanation for the emergence of collective traits in Humans (Boyd and Richerson 2002; Binmore 2011; Smaldino 2014).

In this Chapter, we show that the optimization of collective actions can be reached thanks to individual selection alone. To that end, we take the example of collective hunting. Because it is necessary to agree on which prey to hunt, this behaviour requires the coordination of several individuals (Alvard and Nolin 2002; Alvard 2003; Drea and Carter 2009). This implies that a single mutant choosing a different prey than that agreed upon would not benefit from this behaviour. Thus several equilibria can evolve. The group can specialise on the optimal prey, but it can also specialise on suboptimal ones. In this case no individual mutation can divert the group from a suboptimal prey. Here, we want to understand how individual selection can switch the group's strategy towards the optimal prey.

To that end, we model collective hunting in evolutionary robotics (Nolfi and Floreano 2000; Doncieux et al. 2015). We use evolutionary robotics in this context as it allows to study the nature of the coordination behaviours evolved which, we will show, turns out to have critical effects on collective optimization. Evolutionary robotics has been previously applied to model the mechanistic constraints at play in the mapping between genotype and phenotype (Mitri et al. 2013; Trianni 2014; Bernard et al. 2016b). We design a setting where a pair of individuals can hunt two different types of prey: a suboptimal one and an optimal one. For the sake of simplicity (and not behavioural realism) we choose to arbitrarily call these two prey respectively boar and stag. The reward from hunting a prey depends on the type of prey hunted and the manner in which it is hunted: solitarily or cooperatively. More precisely, two robotic agents are placed in an arena where they can capture non-moving prey. Both these agents are controlled by artificial neural networks, whose connection weights are evolved. Thus, individuals have to coordinate in order to hunt in a cooperative manner. We design the model so that the individuals have evolved the suboptimal equilibrium and that it is not beneficial for a single individual to switch to a different prey alone. We study under which conditions the transition to the optimal equilibrium could occur.

We show that the transition from the suboptimal to the optimal equilibrium is impossible under simple environmental conditions. However, in a more realistic setting, individuals are expected to be able to coordinate. We reveal that when this is the case, the evolution of coordination also enables the switch toward the optimum. Furthermore, the nature of coordination strategies also happens to change the probability for this switch to occur.

Materials and Methods

Experimental Setup

The environment is constituted of a 800 by 800 units arena with four solid walls. We place a collection of circular agents (with a diameter of 20 units) in this arena. The agents are divided between two categories: predators and prey. Predators can move freely in the environment according to the inputs of their sensors which are comprised of several proximity sensors and a camera. Proximity sensors are evenly distributed around the robotic agent's body and have a range of twice the body's diameter. Each of these proximity sensors gives the agent the distance to any obstacle (walls or other agents) it senses in its direction. In comparison, the camera is placed on the front of the robotic agent and is constituted of 12 rays spread in a 90 degrees cone. Each ray gives the type (hunter, boar or stag) and proximity of the nearest agent in its direction. Each hunter begins the simulation next to the other agent at one side of the arena.

In opposition to the predators, prey are static and stay at their initial position until either the end of the simulation or their capture by a hunter (experiments with moving prey capable of avoidance behaviours did not show significantly different results). For a prey to be successfully captured, a hunter needs to remain in contact of this prey for 800 time steps (out of a total number of 20000 simulation steps). The hunter is then rewarded with the value corresponding to the prey. The prey is finally removed from the environment and replaced at a random position in the arena. We thus ensure that both the number of prey and the ratio of the types of prey are kept constant. We consider that collective hunting happens when both hunters are in contact of the prey at the last time step of its capture. This implies that is not necessary for the two agents to be in contact of the prey during all of the hunting time for cooperation to happen.

Each hunter is controlled by a fully-connected multilayer perceptron with a single hidden layer. The inputs of this neural network are constituted of all the sensory information of the agent and a bias neuron whose value is always 1. 1 input neuron is used for each of the 12 proximity sensors and 3 neurons for each of the 12 rays of the camera (the type of an agent is encoded by 2 binary values and 1 value is used to encode proximity). In consequence, the total number of input neurons amounts to 49. The hidden layer is constituted of 8 neurons while the output layer has 2 neurons, each one of them encoding for the speed of each of the agent's wheels. While the connection weights of the network are evolved, the topology is kept identical. Finally, the mapping function used to compute the outputs is a sigmoid.

Simulating Artificial Evolution

The genome has a varying size. Indeed, to increase neural complexity, each individual can evolve two topologically similar neural networks. At the start of evolution, genomes are initialized with 410 real values sampled uniformly in the range $[0, 1]$. This amounts to the number of connection weights of a single neural network. At each generation of the algorithm, a genome encoding for a single neural network has a probability of 5×10^{-2}

to be duplicated. When a duplication event occurs, the genome then encodes separately for the two neural networks, which means that its size is 820. Each generation, there is also a probability of 5×10^{-3} that a deletion event occurs for a genome which encodes for two neural networks. If this happens, one of the two neural networks encoded by this genome is randomly lost and the size of this genome is reverted back to 410. Here we are not interested in a realistic approach to neural modeling. Rather, we want to ensure that the neural complexity of the individuals is sufficient to evolve more complex coordination behaviours.

Evolution is simulated by a classical evolutionary algorithm. At each generation of the algorithm, every individual is paired 5 times with a randomly chosen partner (which may be different each time). Each of these pairs is then independently evaluated in the environment during 20000 steps. In order to decrease the stochasticity effects occurring because of the random positioning of the prey, each pair is evaluated 5 times. Hence 25 simulations are performed for each individual per generation. The fitness of an individual is then computed as the average reward obtained over these 25 simulations.

Selection is conducted according to a (10 + 10) elitist selection strategy. Namely, the population of the next generation is constituted of the 10 best individuals of the previous generation and 10 offsprings sampled from these 10 best individuals. Each offspring is a mutated clone of its parent and no recombination is used. 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} . Finally, each experiment was conducted during 9000 generations and replicated 30 times.

Results

The transition to the optimum is impossible

In this first experiment, we are interested in the transition to the optimum strategy in a very simple environment w.r.t. collective hunting. More precisely, individuals are evolved in an environment constituted of only a single prey of each type. Rewards for hunting are summarized in Table 4.1. The agents are initially pre-evolved during 3000 generations in an environment where hunting stags grants nothing. We thus ensure that the suboptimal equilibrium (i.e. hunting boars cooperatively) has been evolved. Then the individuals are evolved during 6000 additional generations where hunting stags is rewarded.

Prey		Food Reward
Boar	<i>alone</i>	50
	<i>coop.</i>	125
Stag	<i>alone</i>	0
	<i>coop.</i>	250

Table 4.1 – Food Rewards for hunting. Rewards depend on whether the hunt was solitary or cooperative.

We show in Figure 4.1(A) the mean proportion of each type of prey hunted over evolutionary time in 30 independent replications. Figure 4.1(B) shows the proportion of prey hunted at the last generation of evolution in each replicate. We observe that during the pre-evolution step, a cooperative strategy evolved on the boars. However even after this step, there was no transition towards hunting stags. This implies that in this case the switch from the suboptimal equilibrium to the optimal equilibrium is impossible.

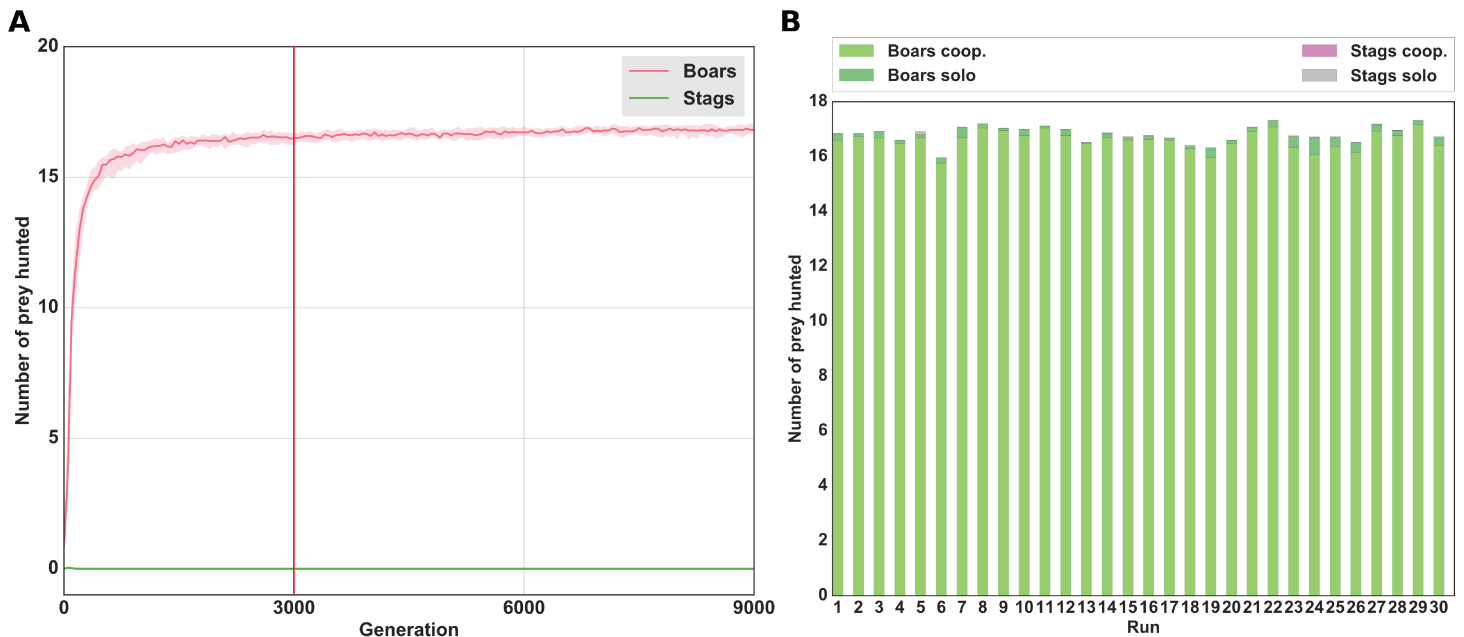


Figure 4.1 – Mean proportion of prey hunted cooperatively. (A) Mean proportion of boars and stags hunted cooperatively by the best individual in each of the 30 independent replications. The colored areas around the medians represent the first and third quartiles. The red line represents the separation between the pre-evolution step (when hunting stags rewards nothing) and the rest of the evolution. (B) Repartition of the prey hunted at the last generation of evolution by the best individual in each replication. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1).

The optimum can be reached under more realistic environmental features

However, it is unrealistic to consider collective hunting only under such extreme ecological conditions. In particular, in nature hunters are expected to collectively choose which prey to hunt from multiple similar prey. To that end, individuals are now evolved in an environment constituted of 18 prey, half of them being boars, the other half being stags. As previously, individuals are first pre-evolved during 3000 generations in an environment where stags do not provide any reward. We then want to study the impact of these new environmental conditions on the transition to the collective optimum (i.e stag hunting).

We show in Figure 4.2(A) the number of replications out of 30 where the optimal equilibrium evolved. More precisely, we consider that this equilibrium was achieved when more than 50% of the prey hunted were stags hunted cooperatively. We present the results from two different settings: the *Control* setting and the *Coordination* setting. The Control setting corresponds to what was presented in the previous Section, where the environment is only constituted of one boar and one stag, to act as comparison with the new results. The Coordination setting corresponds to the setting presented in the last paragraph (i.e. 18 prey in the environment). We observe that the evolution of cooperation on the stags is facilitated as it evolves in 8 replications out of 30.

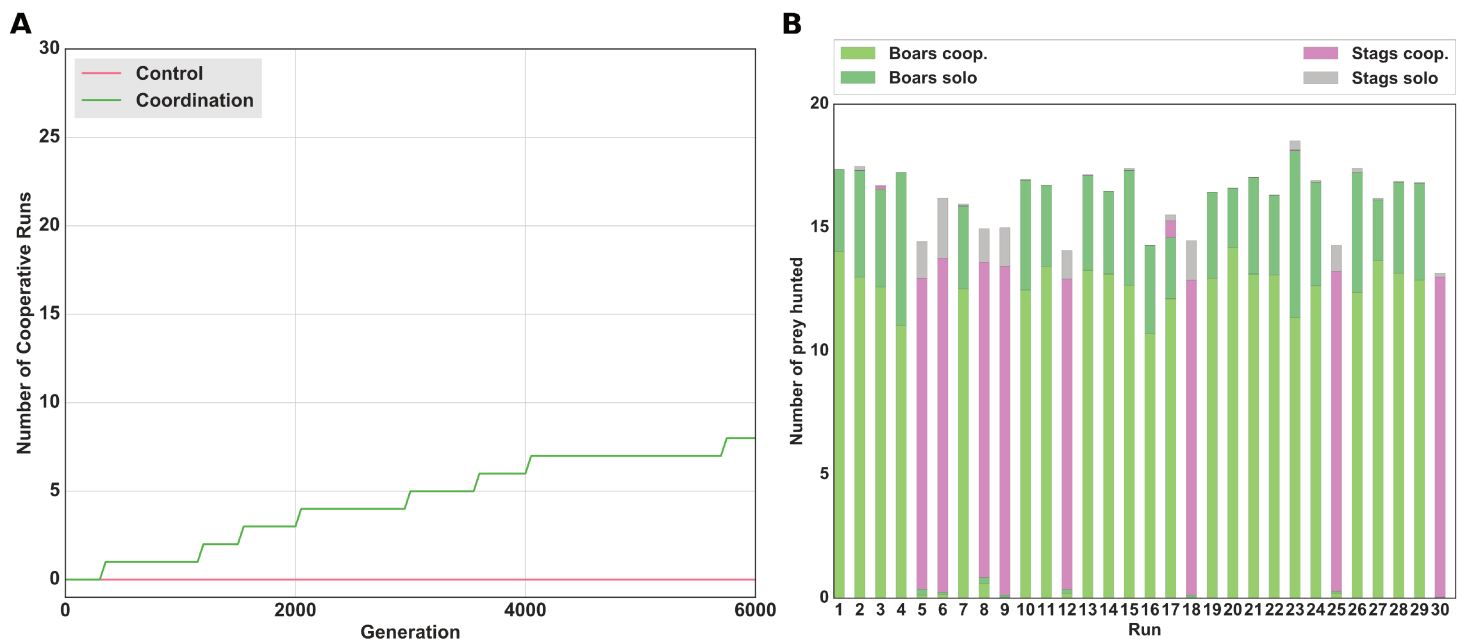


Figure 4.2 – Proportion of stag hunting runs and proportion of prey hunted. (A) Number of replications (out of a total of 30) where stag hunting evolved in the Control and Coordination settings. We consider that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the Control setting, the environment is constituted of one boar and one stag. In comparison, in the Coordination setting, 18 prey are present in the environment and it is thus necessary to coordinate for cooperation to happen. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1). (B) Repartition of prey hunted at the last generation of evolution by the best individual in every replication in the Coordination setting. The population for each replication previously evolved in an environment where hunting stags rewarded nothing.

A striking difference with the previous experiment is that, because of more complex environmental conditions, individuals now evolve a coordination strategy. This behaviour, which we call the *turning* strategy, is shown in Figure 4.3. When individuals adopt this behaviour, they constantly turn around one another. This way, they ensure that they keep the other individual in their line of sight. At the same time, they move towards a prey and, thanks to their proximity, when one of them gets on a prey the other can join it

quickly. In consequence, they are able to achieve cooperation to solve the issue raised by the necessity to decide on which prey to hunt (Figure 4.2(B)). More importantly, because the hunters now react to each other, the behaviour of a mutant can affect that of the other individual. This explains why the transition to the optimal equilibrium was enabled.

However we can observe several drawbacks from this strategy. All these drawbacks come from the fact that both individuals adopt a symmetrical behaviour. In particular, they both steer towards a desired prey and thus react only weakly to the other individual's behaviour. First, as both individuals can guide the group towards a different prey, this may lead to a situation where it is hard to achieve consensus between them. This can considerably slow them down and thus cause suboptimal performance w.r.t. the number of prey hunted. Then, if prey are close to one another, the fact that both individuals can choose a different prey can lead to non-cooperative hunts. Finally, even when the two hunters are moving to the same prey, their constant turning motion to see the other individual implies that they do not take the most direct course towards the prey. Therefore, the evolution of this symmetrical behaviour seems to leave much room for improvement.

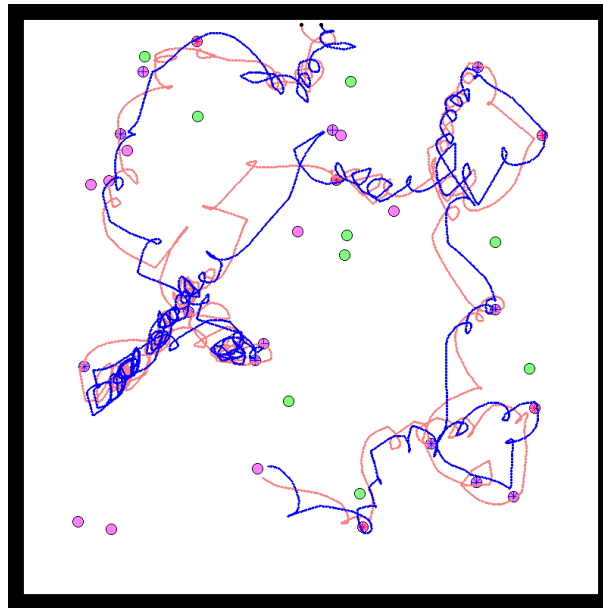


Figure 4.3 – *Display of a turning strategy after an entire simulation. Both individuals adopt a turning strategy during a complete simulation. The paths of the agents are represented in red and blue, starting from their initial positions (represented by black dots). Each disc represents a prey in the 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.*

A more efficient coordination strategy increases the probability to switch to stag hunting

In a next experiment, we address how the evolution of a different coordination strategy could change the outcome of evolution. More precisely, we are interested in the impact of evolving a less symmetrical strategy. We investigate the behaviours of individuals specialised in different roles. To that end, an individual may now duplicate its neural network and thus co-evolve two networks as its controller. The network adopted as controller during the simulation is randomly selected.

Figure 4.4 shows a comparison of the number of replications (out of 30 in each setting) where the best individual evolved a cooperative strategy on the stags in different settings. The *Control* and *Coordination* settings are defined as in the previous Section and serve as comparison. In the *Coordination+Duplication* setting, the environment is constituted of 18 prey as in the *Coordination* setting but the duplication of neural networks is possible. We observe that when increasing neural complexity, the transition to the optimum is facilitated as it evolves in 24 replications out of 30. More importantly, the proportion of replications where the switch to stag hunting occurred is significantly higher than without duplication.

There is a drastic difference between the behaviours evolved with and without two networks. Whereas without the duplication of neural networks we revealed that the individuals adopted a turning behaviour, a *leader/follower* strategy systematically evolved when duplication occurred (Figure 4.5). When agents adopt this strategy, they divide between two very different roles. The *leader* looks for prey, gets on them first and checks but rarely on its partner. In comparison, the *follower* tries to keep the leader in its line of sight at all time and join its partner on a common prey. Therefore, in comparison to the coordination strategy previously presented, the individuals adopt asymmetrical behaviours.

This second strategy is more efficient than the turning strategy w.r.t. the rewards obtained (as shown in Figure 4.6, Mann-Whitney U test on the mean reward at last generation, p -value <0.001). This can be explained by two main factors. Firstly the hunters are faster, as the decision about which prey to hunt is made by only one of the two individuals. Thus they hunt a significantly higher number of prey. Secondly, the asymmetry in the decision making process implies that they are more precise to hunt. Therefore, the proportion of successful cooperative hunts is significantly greater. Consequently, the leader/follower strategy is a collective behaviour that is both more efficient at hunting cooperatively but also leads to a higher probability to evolve stag hunting.

Similar results are observed when roles are not set

We previously showed that the evolution of a leader/follower strategy leads to a more frequent selection of the optimum than with a less efficient strategy. We claim that our results are caused by the general asymmetry in the leader/follower behaviour. Indeed this implies that only one of the two individuals is responsible for taking the decision of

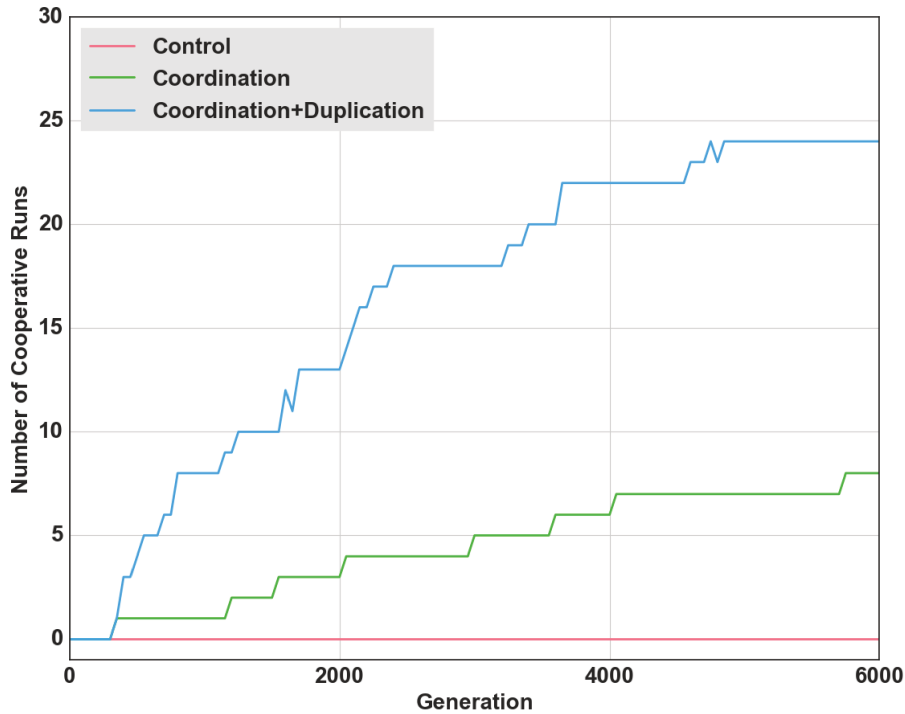


Figure 4.4 – Proportion of cooperative runs. Number of replications (out of a total of 30) where stag hunting evolved in the Control, Coordination and Coordination+Duplication settings. We consider that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the Control setting, the environment is constituted of one boar and one stag. In comparison, in the Coordination setting, 18 prey are present in the environment and it is thus necessary to coordinate for cooperation to happen. Finally, in the Coordination+Duplication setting, the environment is also constituted of 18 prey but individuals have added neural plasticity which allows them to adopt more complex strategies and in particular a leader/follower strategy. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1). The population of each replication previously evolved in an environment where hunting stags rewards nothing.

which prey to hunt. More precisely, the fact that the two individuals are able to use some asymmetrical cue (i.e. the difference in their neural network) allowed them to decide more efficiently. However, as a hunter chooses which network to use at the start of the simulation, this implies that the individuals keep the same role during a whole evaluation. We want to study how removing this constraint would impact the results. To that end, we conducted another experiment where both individuals now begin the simulation with the same neural network. They switch to their second neural network only on a specific event during the simulation. In particular, when one of the hunters gets on a prey, the second hunter switches to its second neural network if possible. This is akin to the first hunter signaling that it found a prey to the latter. When the prey is killed, both hunters switch back to their first neural network.

As in the previous experiment, we compare a *Control* and *Coordination* setting with

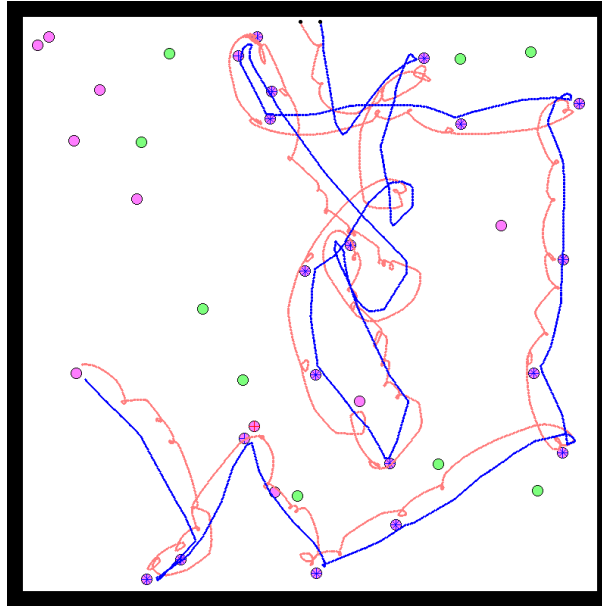


Figure 4.5 – *Display of a leader/follower strategy after an entire simulation.* Both individuals adopt a leader/follower strategy during a complete simulation. The paths of the agents are represented in red and blue, starting from their initial positions (represented by black dots). Each disc represents a prey in the environment. Boars are represented in blue 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.

the *Coordination+Duplication* setting. Only this time, the choice between which neural network to use is done as previously explained. Results are shown in Figure 4.7. We see again that the transition to the optimum is facilitated in the *Coordination+Duplication* setting as in 23 out of 30 replications, there was a transition to the optimal equilibrium. Interestingly, the number of replications where the transition happened is similar to that of the previous experiment. And again, the number of replications where optimization occurred is significantly higher with increase neural complexity than without.

In terms of behaviours evolved, we observe the evolution of another similar asymmetrical behaviour, which we call the *search/join strategy*. More precisely, as long as both individuals use the same neural network, they each look for a prey. However, as soon as one of them finds a prey, the switch in networks for the second individual leads to the adoption of a very different behaviour. Even if the collective behaviour is somewhat different from what could be observed in a leader/follower strategy, this strategy is still more efficient than in the turning strategy (Figure 4.8, Mann-Whitney U test on the mean reward at last generation, p -value <0.001).

Discussion

Because collective hunting is dependent on the simultaneous actions of several individuals, its optimization is difficult. In particular, when a suboptimal equilibrium already evolved,

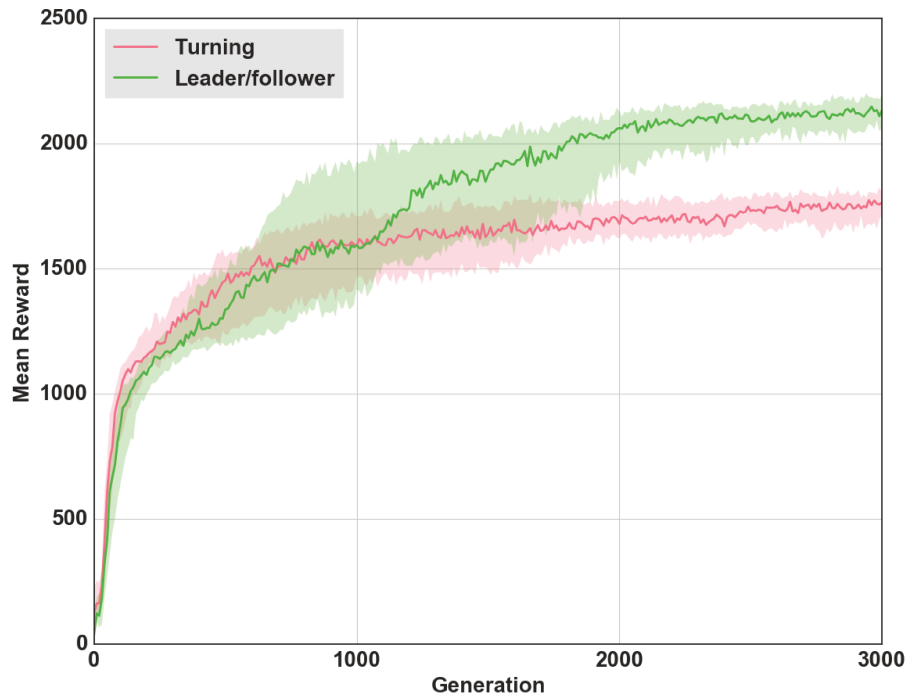


Figure 4.6 – Mean reward comparison of turning and leader/follower strategies. Mean reward of the best individuals in the 30 replications over evolutionary time by individuals adopting a turning strategy (evolved in the Coordination setting) or a leader/follower strategy (evolved in the Coordination+Duplication strategy) during the pre-evolution step. Rewards were 50 for a boar if hunted alone, 125 if hunted cooperatively, 0 for a stag hunted alone and 250 if hunted in a cooperative fashion (Table 4.1).

it is not well understood how the group could switch to the optimum through individual adaptation. Here we showed how the nature of coordination strategies could influence this transition. In particular, our simulations in evolutionary robotics revealed that the recycling of a previously evolved coordination strategy could facilitate the transition to the optimum. In a simple environment where it was not necessary to coordinate for cooperation to happen, the switch to the optimal equilibrium was impossible. Surprisingly, when the environment is more complex and hunting cooperatively is thus more challenging, we revealed that the transition to the optimum was possible. In particular, when coordination was needed to cooperate, the transition to stag hunting was enabled in a significant number of replications (8 out of 30). Indeed, the emergence of stag hunting implies that both individuals change their hunting preference. When individuals coordinate, each one react to the behaviour of the other. This means that a change in the hunting preference of one of the two individuals may sometimes lead to a change in the group’s preference.

But the nature of the coordination strategies evolved is also of utmost importance. To highlight this point, we increased the complexity of the neural networks controlling the agents. This was done by allowing the random duplication of neural networks during evolution. We then showed that, depending on the way the choice of network was done, the individuals would evolve two different strategies: a *leader/follower* strategy and a

search/join strategy. In both of these strategies, the individuals adopt asymmetrical behaviours. We show that when one of these two strategies is evolved, the probability for the transition to the optimal equilibrium to happen is significantly higher (respectively 24 and 23 replications out of 30). This drastic difference can be explained by the asymmetry of both these behaviours. More precisely, the decision making process of choosing the prey on which to cooperate is now one-sided. In the leader/follower strategy, only the leader chooses the prey. In comparison, in the search/join strategy, as soon as one of the two individuals gets on a prey, the other tries immediately to join it. This means that, in both cases, the hunting preference of the follower (whether its role is temporary or not) does not matter. In comparison to the previous experiment now the follower only reacts to the leader. This implies that a change in the leader's behaviour indirectly changes that of the follower. Consequently, a modification in the hunting preference of the individual choosing the prey is sufficient for cooperation on the stag to occur. The mutational distance is thus decreased by both the recycling *and* the evolution of an asymmetrical coordination strategy. From this, it stems that the probability to switch to stag hunting is higher.

More generally, what we reveal through these results is something more critical about the optimization of collective traits. Thanks to the asymmetry of both these strategies, coordination switched from a collective decision making problem to an individual decision making problem. Initially, the transition to the optimum is a collective problem but it now depends on the mutations of a single individual. Furthermore, both strategies are more efficient than the turning strategy which means they can evolve because they are beneficial to the individuals: they are *individually* adaptive. In consequence, individual selection can lead to the adaptation of a collective trait. This suggests that the emergence of other group behaviours may also be explained by such mechanisms, opening a wide range of interesting perspectives on this matter.

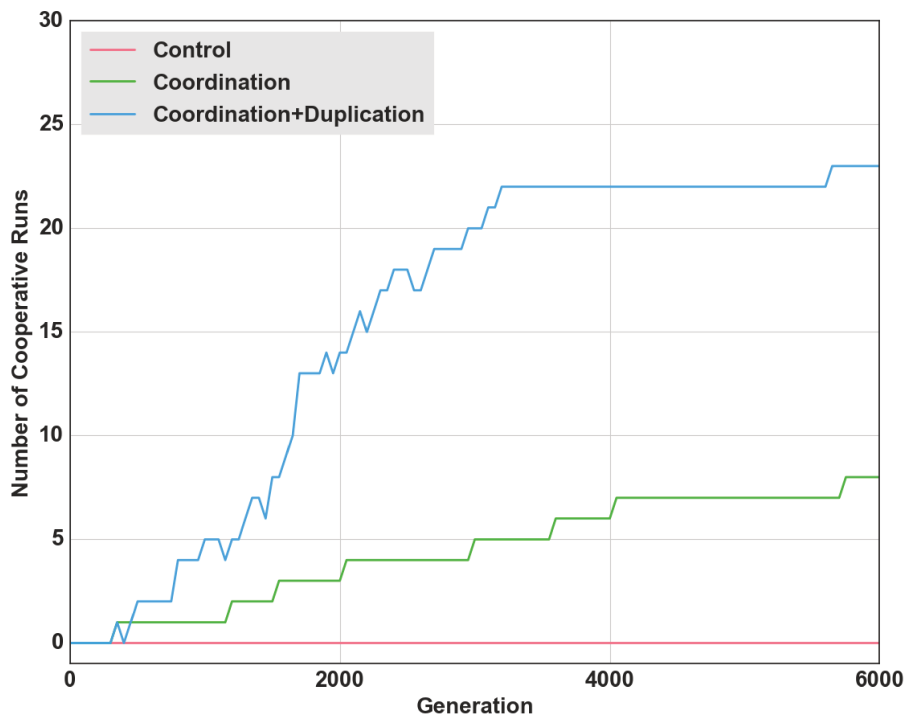


Figure 4.7 – Proportion of cooperative runs. Number of replications (out of a total of 30) where stag hunting evolved in the Control, Coordination and Coordination+Duplication settings. We consider that stag hunting evolved when more than 50% of the prey hunted were stags hunted cooperatively. In the Control setting, the environment is constituted of one boar and one stag. In comparison, in the Coordination setting, 18 prey are present in the environment and it is thus necessary to coordinate for cooperation to happen. Finally, in the Coordination+Duplication setting, the environment is also constituted of 18 prey but individuals have added neural plasticity which allows them to adopt more complex strategies. At a beginning of a simulation, both individuals use the same neural network. A change of network occurs for an individual when the other individual gets on a prey. Rewards for a boar are 50 if hunted alone and 125 if hunted cooperatively. A stag hunted alone rewards 0 and 250 if hunted in a cooperative fashion (Table 4.1). The population of each replication previously evolved in an environment where hunting stags rewards nothing.

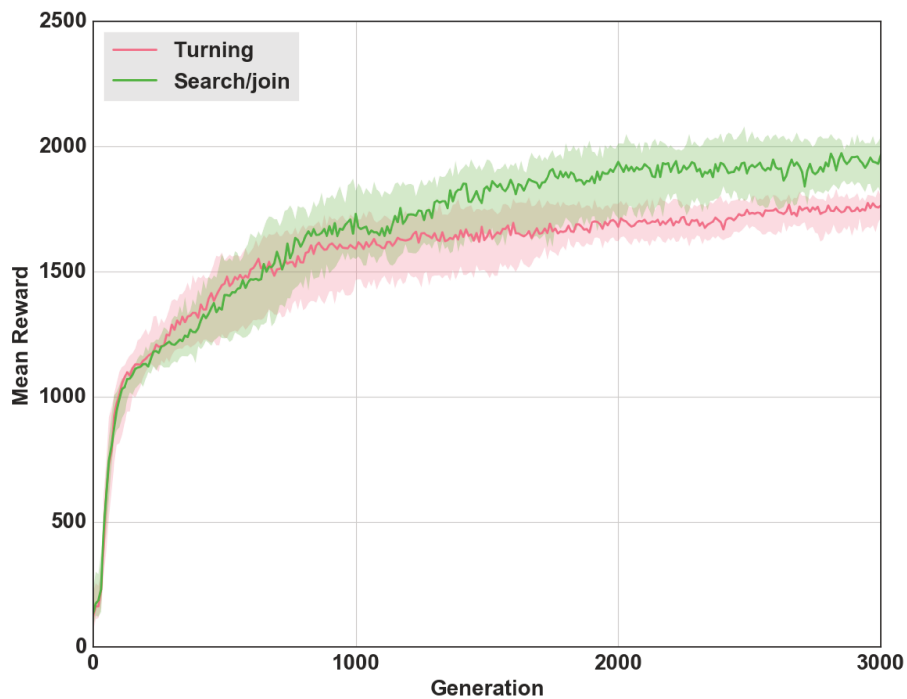


Figure 4.8 – *Mean reward comparison of turning and search/join strategies.* Mean reward of the best individuals in the 30 replications over evolutionary time by individuals adopting a turning strategy (evolved in the Coordination setting) or a search/join strategy (evolved in the Coordination+Duplication strategy) during the pre-evolution step. Rewards were 50 for a boar if hunted alone, 125 if hunted cooperatively, 0 for a stag hunted alone and 250 if hunted in a cooperative fashion (Table 4.1).

Part II

Designing Cooperative Robots in Evolutionary Robotics

Chapter 5

Multirobot Systems and Automatic Design in Evolutionary Robotics

5.1	Multirobot Systems	90
5.1.1	General Properties	90
5.1.2	Architecture Choices	92
5.2	Designing the Control of Collective Robots	94
5.2.1	The Deliberative Approach	94
5.2.2	Manual Design	95
5.2.3	Learning for Automatic Design	96
5.3	Evolutionary Design for Distributed Robotics	98
5.3.1	Evolutionary Robotics	98
5.3.2	Evolving Collective Robots	100
5.4	Genetic Team Composition and the Evolution of Cooperation	101
5.4.1	Team Composition and Levels of Selection	101
5.4.2	Team Composition in Evolutionary Robotics	103
5.5	Conclusions	105

In this Part of the manuscript, we study the automatic design of a distributed multirobot system. In particular, we are interested in the use of evolutionary robotics to design cooperative robots and the influence of genetic team composition (Waibel et al. 2009) in the emergence of efficient coordination behaviours. Namely, we want to study the impact of genetically different robots (i.e. asexual approaches) (Quinn 2001) on (1) the probability to evolve cooperation and (2) the efficiency of the cooperative solutions.

Multirobot systems are many and widely diverse in the way they are designed. Multiple methods have been proposed for both the manual and automatic design of these systems. In this Chapter, we thus review the features (both negative and positive) of evolutionary robotics with regard to the design of distributed robotics in comparison to other techniques. First we give a quick overview of multirobot systems as well as their main advantages when compared to single robots. We emphasize on the design choices that come with building those systems. Additionally, we present a few applications of

multirobot systems that are either seminal and/or noteworthy. Then we focus on the control of collective robotics systems. As there has been a strong interest in taking inspiration from single robots to apply the same solutions to multiple robots, we reveal the particular challenges brought up by these systems. We thus discuss the different manners in which multirobot systems have been manually designed with handcrafted behaviours. Then we address the automatic design of distributed robots thanks to machine learning and reinforcement learning in particular. Our goal is thus to shed some light on the advantages and limits of such approaches. We then move to evolutionary techniques and how they have been applied to the design of multirobot systems. We thus expose the main differences with classical machine learning in this context and quickly review the main results obtained in the design of collective behaviours. Finally, we present the open issue raised by choosing a clonal or a clonal team composition when designing cooperative robots.

5.1 Multirobot Systems

5.1.1 General Properties

Multirobot systems (MRS), or sometimes multi-agent robotics (Dudek et al. 2002), essentially gained fame during the 1980s. The main motivation was to use cooperation between autonomous robots in order to cope with tasks that a classical single robot would not achieve. Multiple advantages may be obtained by using MRS (Cao et al. 1997; Arkin 1998) :

- The parallel execution of multiple robots allows the task to be achieved faster.
- Using multiple robots can ensure robustness and reliability through *redundancy*.
- It can be both cheaper and simpler to produce several simple robots compared to a single complex one (especially if the robot may suffer damages).
- It may be necessary to distribute several robots at the same time to complete the task, in which case a single robot would simply not be sufficient.

This implies that there are several crucial properties that are expected of MRS (Parker 1994). First, MRS are supposed to be *adaptable*. This means that each robot is expected to react to environmental change and, most importantly, to a change in others or induced by them. This also means that, in the lesser decentralized systems, the control system should change the global organization accordingly. Then, a MRS should be *robust*. This implies that the system should not be critically impacted by failure (and in particular individual failures). This is easier said than done but this is also one of the main advantages of such an approach. Because we rely on multiple agents, it is possible to design the system so that a fault on one or several robots does not critically impact the whole system. Finally, it is often expected of MRS to be fully *autonomous*. This means that the system as well as all the agents that compose it should be able to act without human intervention. In

particular, the system should be able to face the unexpected without human control for some time.

From these properties, it stems that there is a range of tasks that are especially appropriate for MRS. While these tasks often relate to real-world applications, they mainly represent general domains in order to design proof-of-concepts on a particular aspect of a MRS (Parker 2000). We can briefly draw a list of the main tasks on which MRS are studied (Cao et al. 1997; Parker 2000; Farinelli et al. 2004) :

Foraging In this sort of tasks, the goal is to collect objects which are scattered in the environment. They then may or may not have to bring these back to a "home". Foraging refers to the real-world tasks of harvesting, toxic waste cleanup or search & rescue. Most often, robots perform the task in a very independent manner, where the individual behaviour of a single entity does not really impact that of others. However, they may also rely on communication and *stigmergy* in particular, i.e. indirect communication achieved by the previous modification of the local environment by an individual. Additionally, foraging tasks also have strong ties with biology and the behaviours of eusocial insects in particular. A main challenge in foraging is for the robots to efficiently explore the environment, i.e. without repeating each other's actions.

Collective transport The goal here is for several robots to collectively push an object (also called box pushing). This object is usually too big or too heavy for a single robot to move it alone and it thus requires the coordination of several individuals. A specific type of box pushing, cooperative manipulation, may require robots to carry the objects to a destination (rather than pushing). Box pushing may not necessarily require that robots be aware of the others for the task to be achieved (Sen et al. 1994).

Collective motion A popular task is to design robots that are able to move in a coordinated manner. This may imply that we simply desire robots to move together towards a path, as in the case of flocking, or that we want them to adopt a particular formation for the duration of the motion. This sort of task can often be accomplished by agents with minimal capabilities in terms of sensors, effectors and communication. One of the central issues studied in collective motion is the design of simple and local (i.e. individual-level) control rules that allow for the collective emergence of the desired behaviour.

Traffic control Another common task is that of traffic control. This is a problem of multirobot path planning, where several individuals often have their own personal goal. They must then coordinate with others in order to accomplish their goal without causing collisions or deadlocks. This is akin to a problem of resource conflict where robots have to share the environment with others.

Monitoring Monitoring refers to the task of using multiple robots to observe and track a defined number of targets moving in the environment. Robots have thus to cooperate

in order to ensure that all targets are monitored during the longest amount of time. In particular, there is a strong emphasis on coordination so that the agents can efficiently follow the targets and switch between them when necessary.

5.1.2 Architecture Choices

Control	Centralized	Decentralized	
Team composition	Homogeneous	Heterogeneous	
Communication	Environmental	Passive	Intentional
Group size	Small (~ 10)	Swarm (> 100)	

Table 5.1 – *Architecture choices in multirobot systems.*

Because applications vary greatly, there is no canonical architecture for MRS (Cao et al. 1997; Parker 2008). There are some more popular choices that we will highlight here but mainly one can use what best fits his/her needs. The main design choices w.r.t. architecture are the following (see also Table 5.1 for a quick summary) :

Control The control of an MRS can be *centralized* or *decentralized*. In a centralized architecture, a single agent is responsible for controlling the system. Thus while this agent has full knowledge of the whole system, it represents a critical point for failures. Therefore, this type of organization is rare in MRS and most use a decentralized approach (Parker 2008). However, the work of D’Andrea (D’Andrea 2012) on the development of the Kiva systems, where a large group of robots (hundreds) move in a warehouse to bring products to the workers, is of note in this category. In particular, a central control was responsible for the coordination of all the robots. In comparison, decentralized architectures can be of two types: *hierarchical* or *distributed* (Cao et al. 1997). In a hierarchical architecture, the system is locally centralized and some agents are in charge of a group of other agents to organize the task at hand. For instance, in one of the very first successful MRS, CEBOT (Fukuda et al. 1988), particular robots (called "master cells") could communicate with other master cells and allocate subtasks to all of the agents in the system. On the contrary, in a distributed system, all agents are equal w.r.t. control which, while robust, implies that it is harder to achieve coherence between every agent.

Team composition It is possible to use *homogeneous* or *heterogeneous* groups of robots. In an homogeneous team, individuals are all identical in terms of both software (control) or hardware (morphology and sensors). In comparison, heterogeneous robots vary between one another on any or both of these aspects. In consequence, homogeneous teams are more resilient to failures as every agent has the same capabilities, thus decreasing the impact on the system of losing a given individual. It is also easier to allocate tasks between robots because every agent can perform equally. However heterogeneous groups allow to benefit from differences between individuals to achieve more diverse behaviour, in particular when coordination is required.

Communication We can mainly divide the type of communication implemented in MRS into three categories: *environmental*, *passive* and *intentional* (Cao et al. 1997; Parker 2008). Environmental communication refers to the indirect communication we briefly mentioned previously: stigmergy. Stigmergy means that the agents will sense modifications in the environment done by a previous agent and use this information to modify their behaviour. It is as if one agent indirectly communicated to the other through the environment. This type of communication is thus limited by the capabilities of agents to perceive complex information from the environment. Passive communication is another type of indirect communication where the agents rely on their sensors to observe the actions of others in the group. Thanks to this sensory feedback, robots can interact with each other without needing direct communication. However, it shares the same limitations as with stigmergy. Finally, intentional communication refers to direct communication between the robots. This thus allows to exchange complete information between teammates.

Group size On this point, MRS are really diverse and the number of robots involved in a collective task can scale from two to a thousand (Rubenstein et al. 2014). A smaller group size usually means that it is possible to design more morphologically complicated robots where each individual may have elaborate capabilities. On the contrary, when we are interested in bigger teams, individual capacities tend to decrease in favour of collective complexity. However, an open challenge is to scale up algorithms that were designed for a small group of robots to a larger team. Large groups of robots are often referred to as *swarms* (Beni 2005). In this case, robots in a swarm often possess very basic sensory capabilities and may not achieve much on their own. The emphasis is put on the *emergence* of collective functionalities from individual interactions (Kube and Zhang 1993; Parker 2008). This means that we expect to see the appearance of global collective complexity from the local interactions between agents of the swarm, a process also known as *self-organization*. More precisely, swarm robotics are based on the principle of superadditivity (Parker 2008), where the whole result (collective behaviour) is better than a simple sum of all its parts (the agents' behaviours).

In this manuscript we study a particular instance of MRS. Namely, we focus on the design challenges of fully distributed MRS (i.e. decentralized) where a small group of robots are morphologically homogeneous. In this context, we are interested in the critical differences that come with using a team of homogeneous or heterogeneous robots w.r.t. control. Lastly, as communication is not the focus of our study, we do not explicitly give the robots particular communication capabilities. Therefore, if any communication takes place between individuals then it can be considered to be passive. The rest of this Chapter is to be understood in the context of these specific architecture choices.

5.2 Designing the Control of Collective Robots

So far we have presented a general overview of the properties and choices that come with designing a MRS. Namely, we discussed the *what* and *why* of multirobot systems. In this Section, we focus on the *how*. More precisely, we are interested in the design techniques involved in creating a distributed multirobot system. Designing the control of a robot mainly depends on its situatedness, i.e. the complexity and uncertainty of the environment in which it operates (Matarić and Michaud 2008). Most of the early design techniques used in MRS have been inspired by classical single robots techniques.

5.2.1 The Deliberative Approach

This approach is also referred to as the Sense-Model-Plan-Act architecture (Albus 1991; Iocchi et al. 2001; Matarić and Michaud 2008). This has been the classical approach in robotics (Nilsson 1984) and AI and is concerned with representing high reasoning capacities. As such, it has also historically been one of the first manners in which to approach MRS design. The basic principle is that all sensory information is computed under the internal knowledge of the robot in order to plan and determine the next action. This means that these architectures are based on an internal representation of the world. The model is often constituted of a set of symbols which are computed by a logical system. However, planning is a classical problem in AI and is known to be time costly. Therefore, while this architecture would be the most efficient in a perfect world, the process of building a world representation and planning is computationally expensive and lacks critical real-time reactivity.

In the case of multirobot systems, we may define an additional global level of control as *social deliberative* (Iocchi et al. 2001). In a social deliberative MRS, a global strategy will be planned so that the organization of the whole system (e.g. task allocation) can handle environmental changes. This type of MRS may have a global representation of the world shared between the agents but it is not necessary. Do note that, as we will see, the global control design of the MRS often differs from that of the individuals (e.g. a social deliberative system may be composed of behaviour-based robots).

Mostly, when a deliberative approach is adopted, it is used as a global level control only. Werger and Matarić (Werger and Matarić 2000) designed a MRS tasked with multi-target observation (i.e. monitoring) with a social deliberative distributed MRS. While robotic agents were individually behaviour based (see next Section), group-level deliberation was achieved thanks to an architecture they called the "Broadcast of Local Eligibility". More precisely, each robot could evaluate its eligibility to accomplish a given task and then broadcast this value. The robot with the highest value could then claim the task. They were thus able to solve the issue of task assignment in their system. In the context of the RoboCup (Kitano et al. 1997), a soccer competition between teams of robots, Candea and colleagues (Candea et al. 2001) developed a distributed and heterogeneous soccer team in a deliberative system. In their proposed architecture, each robot was capable of playing any role and they could switch during a match. In particular, a voting system coupled with the possibility to evaluate and communicate the utility of an agent for a

given role was used to efficiently assign these roles to the individuals. They could thus achieve strong coordination between distributed and heterogeneous robotic agents.

5.2.2 Manual Design

Because deliberation alone is hard to achieve in a distributed multirobot context, there has been a strong interest in using ad-hoc methods for the programming of robots behaviours.

Reactive approaches

In light of the complexity of deliberative architectures was born the opposite stance: the *reactive approach* (Brooks 1986). In comparison with deliberation, this architecture is not based on reasoning nor planning. Rather, there is a direct connection from sensors to effectors, inspired by the biological concept of stimulus-response. This architecture is usually constituted of a programmed set of rules which, given the sensory inputs, return the desired output actions. This implies that reactive systems can achieve very fast computation and thus are convenient when quick reaction is necessary. However, as robots do not keep any representation of the world and most often do not store any information, they are basically myopic. This can be useful when *a priori* knowledge of the environment is sufficient but does not fare well with uncertainty and novelty. Additionally, there is a strong emphasis on the concept of emergence we previously mentioned. Because the control of agents under a reactive approach does not allow for complex individual capacities, we expect functionalities to come from the emergent collective behaviour. This means that designing a reactive distributed system is a bottom-up approach which often implies a back and forth between handcrafting individual behaviours and observing the collective result.

A popular way to design reactive behaviours is also to take inspiration from nature. For instance, the well-known work of Reynolds on the "boids" (Reynolds 1987) was inspired by the collective motion of flocks of birds. He developed a simulation of the collective behaviours of a swarm of simple agents with no cognitive capabilities under three ordered basic rules: collision avoidance, velocity matching and flock centering. Namely, agents should avoid collisions with neighbours, match their direction and speed (i.e. velocity) with that of others in the flock and stay close to other individuals. He was able to thus design a group of agents behaving in a similar way as a flock. Hauert et al. (Hauert et al. 2011) implemented Reynolds' boids on 10 real flying robots. In particular, they were interested in the influence of communication range and turn rate on the emergence of flock-like behaviours. They showed in simulation that communication range needed to be high enough so that coherent flocks could be maintained. Rubenstein and colleagues (Rubenstein et al. 2014) used a swarm of a thousand individuals capable of self-assembly. The robots, called kilobots, were embedded with an infrared transmitter and receiver to communicate with neighbours and measure their proximity. The system was given a 2D image by a user defining the shape the robots must replicate. One after the other, the agents then used very simple local rules by following the edge of the group and tracking the distance from their origin to collectively organize into the desired shape.

Behaviour-based approaches

Behaviour-based approaches (Arkin 1998) were proposed after the appearance of reactive approaches, with the desire to improve on the latter's capacity to react to dynamic environments. This was mostly introduced and popularized by Rodney Brooks subsumption architecture (Brooks 1986). In behaviour-based architectures, the robot control is constituted of several basic behaviours, which are organised in separate modules. In a similar way as the reactive approach, these behaviours are directly connected to the sensors and will activate according to a certain set of rules. However, in comparison to a purely reactive approach, these behavioural modules can keep a state as well as a representation of the world, allowing for higher reasoning and planning. Additionally, they are connected to each other in a hierarchical fashion. Those modules are designed to interact with one another in order to collectively achieve the task at hand. Complexity is thus expected to emerge from the interactions between low-level behaviours. In consequence, behaviour-based architectures are efficient when the environment is dynamic but pure reactivity alone is not sufficient. In a similar way as reactive control, these architectures are usually designed in a bottom-up approach where behaviours are coded incrementally as building blocks in an increasing complexity.

Behaviour-based approaches are among the most used for handcrafted robot control in MRS (Arkin 1998; Matarić and Michaud 2008; Parker 2008). For instance, Parker proposed and developed the ALLIANCE architecture which she successfully implemented on real robots (Parker 1994). The issue was to design a fault-resistant system of heterogeneous robots which could achieve coordination. Based on a subsumption architecture, the system was composed of low-level behaviours combined together in a set to accomplish a particular task. Then, given environmental information (e.g. which task is being taken care of, what needs to be done), a motivation was computed to select the appropriate behaviour set. Matarić proposed the *Nerd Herd* (Matarić 1995), a group of 20 identical robots capable only of detecting obstacles and other robots. Each robot was constituted of the same set of pre-programmed behaviours as obstacle avoidance, homing, aggregation, dispersion, following and safe wandering. The system could combine these behaviours in order to achieve higher functionalities. For instance, collective foraging was obtained by applying a temporal combination operator in order to switch between avoidance, dispersion, following, homing and wandering.

5.2.3 Learning for Automatic Design

In comparison to the approaches presented in the previous Section, there has been also a strong interest in using automatic design for MRS. In particular, handcrafted behaviours do not fair well in the face of uncertainty and varying environmental conditions and may require a tedious back and forth before the adequate collective behaviour is obtained. In consequence, there is a large body of work on using learning for the design of MRS. For instance, in an article published in *Science*, Werfel et al. (Werfel et al. 2014) focused on real robots capable of building structures. They used a team of homogeneous robots, taking inspiration from the mount-building capabilities of termites. The goal was to achieve the

construction of a user-specified structure with 3 robots and small identical bricks. While the robots were fully reactive and relied only on stigmergy, the rules they followed were automatically generated. This was done by an offline compilation process which, given the building structure, generated a representation of movement guidelines for robots, akin to traffic laws. They showed that their group of robots was able to collectively construct the structures as well as dynamically react to changes in the built structure.

However, most learning techniques in MRS rely on machine learning. Machine learning has always been a critical challenge in artificial intelligence. Thus it has naturally been applied to robotics (Hertzberg and Chatila 2008). Classical machine learning can be divided into three different categories: supervised, unsupervised and reward-based. While the goal in machine learning is generally to optimize performance (e.g. for classifiers), the emphasis in mobile robotics is that the robot may adapt quickly. As such, most of the literature on learning in robotics has been focused on reward-based techniques (Mataric and Michaud 2008), most commonly referred to as *reinforcement learning* (RL) (Sutton and Barto 1998). RL rests upon the mathematical framework of markov decision processes (MDP) (Bellman 1957). In RL a robot learns an optimal policy (i.e. a sequence of actions depending on the states the robot is in) thanks to a value function. Learning is thus achieved through rewards and punishments attributed to the robot according to its actions. The general goal in RL is to estimate the value function. This value function corresponds to the expected value of a state given a certain policy.

The main RL method applied to robotics is temporal-difference (TD) learning (Sutton 1988; Bradtke et al. 1996). Based on the principles of TD learning, two major algorithms have been developed: on-policy SARSA (State-Action-Reward-State-Action) and off-policy Q-learning (Watkins 1989). Additionally, most RL techniques have theoretical proofs of convergence¹ (Panait and Luke 2005).

In the case of learning for multiple robots, the process is more challenging. In particular, other robots are often expected to be learning at the same time. At the very least, the learning process must take into account the presence of other dynamic agents. Yet the theoretical foundations behind MDPs rest upon the assumption that the environment is stationary (Littman 1994; Parker 2008). Consequently, adapting RL methods to multiple robots is not trivial. However, there is an extensive literature on learning in multiagent systems (MAS), of which some can be applied to MRS (Stone and Veloso 2000; Yang and Gu 2005; Panait and Luke 2005). In particular in the case of distributed control, concurrent learning, where each individual is an independent learner, has been widely studied. To that end, the framework of Dec-POMDP (for Decentralized Partially Observable Markov Decision Process) has been of interest in the MAS community (Bernstein et al. 2002; Amato et al. 2013). This model is a decentralized extension of basic POMDP (Åström 1965), which deals with real world partial observability. However, the optimal resolution of a Dec-POMDP (i.e. finding an optimal joint policy) was proven to be NEXP-complete (Bernstein et al. 2002) and thus intractable. As such, most ap-

¹Please note that what we have presented here is only a crude summary of RL in order to give sufficient context to the rest of our discussion. We point those interested by the subject to more exhaustive literature (Sutton and Barto 1998; Deisenroth 2011)

proaches rely on approximations of the model in order to solve the problem (Beynier and Mouaddib 2011; Amato et al. 2013). For instance, Seuken & Zilberstein (Seuken and Zilberstein 2007) have combined top-down heuristics with bottom-up dynamic programming to ensure linear complexity w.r.t. horizon length. Called memory-bounded dynamic programming, this method uses heuristics in order to limit the number of agents policies generated by dynamic programming. Dinbangoye et al. (Dibangoye et al. 2015) assumed that interactions between agents took place locally in order to exploit the separability of the value function. This way they could transform a Dec-POMDP into a MDP to significantly gain in scalability. They thus could solve a task between up to fifteen agents while preserving convergence to an optimal solution.

However, as big as the literature on learning in MAS is, transferring reinforcement learning techniques from MAS to MRS still represents a challenge (Yang and Gu 2005). In particular, while results in MAS offer interesting perspectives, MRS necessitate continuous actions and/or states spaces. This is something which is not that much studied in classical MAS. To overcome these problems, and the issue of continuous spaces in particular, several different solutions based on approximations have been proposed. For example Mataric proposed to extract the features from the learning space by reformulating states and actions into conditions and behaviours (Mataric 1997). This way, the size of the spaces was greatly decreased. She also implemented shaping (i.e. decomposing a complex task into several simpler subtasks which are then learned in succession) in order to ease the learning process. In comparison, Fernández and colleagues (Fernández et al. 2005) developed a learning MRS by discretizing the states space and then applying an algorithm to generalize from this discrete space. This particular algorithm, called ENNC-QL, is based on a supervised approximation of the value function. In the case of an adversarial MRS learning for soccer, Bowling & Veloso (Bowling and Veloso 2003) introduced GraWoLF (for Gradient-based WoLF). They used a policy gradient technique, which was proposed to overcome intractable and continuous states spaces (Sutton et al. 2000) as well as WoLF (Win or Learn Fast), an algorithm to ensure convergence in the context of concurrent learning. Lastly, Stone & Sutton (Stone and Sutton 2001) also proposed a reinforcement learning method for a soccer competition. They implemented a SMDP (Semi-Markov Decision Process) Sarsa(λ) with linear tile-coding function approximation. With this method, they were able to have robots learn in a keepaway task.

5.3 Evolutionary Design for Distributed Robotics

5.3.1 Evolutionary Robotics

In this thesis, we study a different technique for the automatic design of robots: evolutionary robotics (ER) (Nolfi and Floreano 2000; Doncieux et al. 2015). As we already covered some of the more technical details of ER in the Introduction, we are more interested here in briefly reviewing the contributions to the field of distributed MRS and highlight the challenges that are associated with. The key idea behind ER is to apply concepts of evolutionary computation to the design of robots. This means implementing concepts

of selection and variation to build robust and adaptable robots. We previously showed that several design techniques have often been inspired in part by biology. For example reactive controllers are inspired by the concept of stimulus-response (Brooks 1986), swarm robotics took inspiration from the collective behaviours of eusocial insects (Bonabeau et al. 1999) and it is sometimes argued that major advancements in reinforcement learning mimic natural cognitive processes (Montague et al. 1996). Thus there is an interest in taking inspiration from evolution for the design of complex machines. In particular, ER uses evolution to approach robotic design in an holistic manner. The robot is considered as a whole and the evolution of its behaviour results from the interactions with the environment (and the other individuals in the case of MRS): ER works on embodied agents. Additionally, the evolutionary process is used as a meta-heuristic to search through the space of candidate solutions. In particular, ER works well in open environment because it can be used as a black-box optimization technique thanks to a loose formulation of the objective function. As such a minimum set of assumptions have to be made when using ER to design a robot (Bongard 2013). The major open issues in ER are on the transferability to real robots (i.e. reality gap) (Mouret et al. 2012; Cully et al. 2015), the genotype and phenotype encodings (e.g. evolution of neural network topology (Stanley and Miikkulainen 2002)) and the selective pressures applied by the evolutionary algorithm (Lehman and Stanley 2011; Mouret and Doncieux 2012).

At its core, ER can be considered as a learning technique. However, it may be troublesome to classify ER among learning algorithms. Indeed, while ER is a learning process in the machine learning sense of the word (i.e. a process which improves and optimizes candidate solutions according to a certain goal), it is not the case in a more biological sense: evolution is a phylogenetic adaptation while learning is an ontogenetic adaptation. This difference is even more critical now that combining evolution and learning represents an open issue in the field (Urzelai and Floreano 2001; Mouret and Tonelli 2014; Doncieux et al. 2015). Here we thus are careful to use the latter (i.e. biological) definition of learning and refer more precisely to reinforcement learning in this case. It is also important to note that ER and RL share several similarities (Whiteson 2012; Stulp and Sigaud 2013; Doncieux et al. 2015). In both frameworks, the goal is for a robot to evolve (or learn) a behaviour (which is akin to a policy in RL) which maximizes a particular value: rewards in RL or fitness in ER. In particular, we can compare ER to a direct policy search in RL (Kober et al. 2013) because it does not focus on finding an estimation of the value function of the states and actions but exploits the global value (i.e. fitness) of a policy. However, ER often necessitates higher computational time to find a good solution while dynamic programming in RL is guaranteed to find an optimal policy in polynomial time (Littman 1994; Whiteson 2012). In comparison ER works very well under partial observability and with problems that would require continuous or a large number of states in RL (as is the case for distributed MRS). In particular, ER explores the space of behaviours rather than that of states (Panait and Luke 2005).

ER can also be used as an online design method. In the "classical" framework of evolutionary robotics, the evolutionary algorithm is called offline. This implies that there are two distinct phases in the development of a robot: the design phase (i.e. evolution of controllers) and the operational phase (i.e. deployment of robots) (Doncieux et al.

2015; Francesca and Birattari 2016). This thus is based on the assumption that the environment where the robots are deployed is the same that the one where they were evolved. Or at least it considers that the evolved controller will be capable of adapting to the new environmental conditions. In an online method the design process is done directly in the operation environment. In the case of multiple robots, this gave rise to distributed online evolutionary robotics, often called *embodied evolution* (Ficici et al. 1999; Watson et al. 2002). Because of the complexity of learning exact policies with multiple robots, the field of online evolution has sparked greater interest in multirobot settings than with a single robot (Doncieux et al. 2015).

5.3.2 Evolving Collective Robots

While ER has been mainly focused on the design of single robots (Nolfi and Floreano 2000; Doncieux et al. 2015)², its potential for the engineering of complex collective systems is well known (Baldassarre et al. 2003b). For instance, Reynolds (Reynolds 1992) proposed an evolved version of its "boids" simulation. He evolved a herd of between 16 and 20 critters whose goal was to avoid both obstacles and a predator. He used genetic programming where the evolved programs (i.e. genotypes) were Lisp expressions based on simple behavioural functions: turn, look-for-obstacle, look-for-friend and look-for-predator. He showed that he could evolve vision-based coordinated motion for a herd that avoided collisions.

More generally, evolution has been widely used in the context of swarms (Brambilla et al. 2012; Francesca and Birattari 2016). In particular, it allows to divert from the classical approach of manually designing individual behavioural rules. On the contrary, ER can automatically evolve self-organized control according to a group-level fitness score. Evolution was used in the context of the Swarm-bot project (Mondada et al. 2005), where the goal was to engineer a swarm of simple identical robots capable of using self-assembly to navigate across rough terrain and achieve different collective tasks. In particular, Baldassarre and colleagues (Baldassarre et al. 2003a, 2007) achieved coordinated motion between a swarm of 36 simulated robots. Because the robots were connected to each other in line, they had to coordinate their movement in order for the whole swarm to reach the objective. Robots were controlled by a neural network which, given the traction on the robot in 4 directions, computed the desired motion. They showed that the evolved robots were also capable of high adaptability and generalization under various environmental conditions: number of robots, shape of the swarm, variation in the rigidity of the robots' connections, rough terrain and robots connected through a passive object. The controllers were also successfully transferred on real robots. The swarm-bot framework was also investigated for collective transport of objects by Groß and Dorigo (Groß and Dorigo 2004). They evolved the neural networks of up to 16 autonomous robots that were capable of collectively push or pull an object towards a moving target. Trianni et al. evolved a swarm of robots displaying aggregation behaviours (Trianni et al. 2003) and that later

²As the focus of this manuscript is on multirobot systems, we will not discuss the literature on the subject of ER for single robots. Interested readers should direct their attention towards more extensive reviews of the field (Floreano et al. 2008; Bongard 2013; Trianni 2014; Doncieux et al. 2015).

cooperatively navigated an environment to overcome their limited sensory capabilities and avoid falling into holes (Trianni et al. 2004). In the context of swarms, Duarte and colleagues (Duarte et al. 2016) developed 10 simple and small real robotic boats that they evolved for basic tasks: homing, dispersion, clustering and area monitoring. They used artificial neural networks evolved with NEAT (Stanley and Miikkulainen 2002). The best 3 controllers of each evolutionary run were then transferred into real robots to test them. Finally, Hauert et al. (Hauert et al. 2009) evolved a group of 20 simulated flying robots, or MAVs (Micro Air Vehicles), tasked with the establishment of a communication network. Launched from a human rescuer, the robots had to coordinate to find the other rescuer and then set and maintain a multi-hop communication link between these two rescuers. Every robot was controlled by a neural network which outputted the turn rate of the MAV given the heading compass of the robot and the number of network hops that separate it from the two rescuers. The connection weights of the network were encoded in a binary string and evolved by a genetic algorithm.

Additionally, D’Ambrosio and Stanley have been interested in using HyperNEAT (Stanley et al. 2009) to evolve teams of agents capable of coordination. For instance, they implemented (D’Ambrosio and Stanley 2008) HyperNEAT in a predator-prey experiment. In particular, they evolved neurocontrollers for predators’ behaviours which had to hunt moving prey. The predators could not see other predators and thus had to learn complementary roles so that they would not interfere with each other. They used Computational Pattern Producing Networks (CPPNs) which allowed for the agents to assign roles depending on their relative geometry. Furthermore, they showed that seeding evolution with the genome of a single pre-evolved agent could benefit the learning process by injecting domain knowledge. D’Ambrosio and colleagues (D’Ambrosio et al. 2012) investigated coordination between 4 robots thanks to direct neural network communication. More precisely, the neural network of each robot was connected to the internal nodes of the networks of other agents, which they called the *hive brain*. Robots could simply move left or right and the task was for them to synchronize their motion (akin to several pendula) with no sensory information about other robots. They showed that the agents could evolve an efficient communication strategy that led them to synchronize their motion. Moreover, the evolved controllers could transfer well to real robots.

5.4 Genetic Team Composition and the Evolution of Cooperation

5.4.1 Team Composition and Levels of Selection

While ER is often viewed as a black-box optimization framework, several critical design decisions impact its efficiency when applied to robotics (Trianni 2014). As such, these design choices represent a specific problem in the field of evolutionary robotics (Doncieux and Mouret 2014). We are interested here in the evolution of cooperative robots in ER. In that context, two features of the evolutionary algorithm are especially critical (Waibel et al. 2009; Lichocki et al. 2013):

Level of selection This represents the level at which selection is applied, which is in part impacted by the way fitness is distributed between individuals.

Team composition This is the genetic composition of the team of robots, which corresponds to the manner in which robots from a group are encoded given the population of evolved genotypes.

Selection can act at the level of the *group* or that of the *individual*. Namely, the level of selection is concerned with the way fitness is attributed to each individual in the team. If group-level (or team-level) selection is used, every individual is equally rewarded by the team's performance in the task. This means that in the case where every individual separately evolves, the evolutionary process may be slowed down by the fact that the performance feedback may not be adequately tailored to the individual. In comparison, under individual-level selection, each individual is rewarded based on its own performance. This in turn means that the emergence of cooperative behaviours is not ensured as the individuals could benefit from selfish actions. This problem is often known in multiagent learning as "credit assignment".

On the other hand, team composition is a well-known design choice for MRS in general (as we previously talked about in Section 5.1). Here we focus solely on the issue of team composition with relation to the control of robots, regardless of morphology. In evolutionary robotics, every individual in an homogeneous team is composed of the same genotype. This approach is thus often called a clonal approach (as agents are clones of each other). In comparison in an heterogeneous group every individual is encoded by a different genotype, a process also known as an aclonal approach.

Homogeneity tends to facilitate the maintenance of novel beneficial mutations that could be lost in an heterogeneous context (Quinn 2001). As such, using homogeneous teams can lead to finding solutions in less computational time. Also, because the performance of one individual (w.r.t. fitness score) is the same as every other agent in the team, it should be easier to evolve cooperative solutions. Indeed, an individual can benefit from its behaviour as soon as it benefits the whole group³. Additionally, because every individual has the same control, it is easier to achieve coordination when it is expected that the agents behave similarly. However heterogeneity implies that individuals are different and as such may rely on diverse capabilities. In particular, when it is expected that the agents work together in complementary ways (e.g. division of labour), this behavioural asymmetry is easier to achieve with heterogeneous individuals. However an issue in heterogeneous teams is that, depending on the level of selection, the evolution of cooperative solutions may be hindered by the exploitation of selfish behaviours. Indeed, if each individual can selfishly benefit from its behaviour, it may not contribute to the collective action. The main approaches that have been adopted in ER have been centered on using team-level selection with homogeneous teams or individual-level selection with heterogeneous teams (Waibel et al. 2009).

³Given the strong ties of this thesis with biology, it is interesting to note that this process is similar to the evolution of altruism under kin selection mechanism (which requires genetic relatedness between individuals).

5.4.2 Team Composition in Evolutionary Robotics

When evolving cooperation is concerned, the classical approach is to use homogeneous teams of individuals (or clonal approaches). In particular, a large part of the literature is concerned with swarms, which most often are constituted of homogeneous individuals. For instance, the works in swarm behaviours presented in the previous Section all evolved a single population of genotypes where each genotype encoded for every agent in the group. As we previously explained, this can be explained by the fact that homogeneous teams are a good choice when evolving cooperation and strongly coordinated behaviours. Waibel and colleagues (Waibel et al. 2009) gave an experimental proof of this assertion in a work dedicated to this issue. They produced a study on the influence of team composition (homogeneous or heterogeneous) and level of selection (team-level or individual-level) in 3 different foraging tasks that did not require specialisation: an individual one, a cooperative one and an altruistic one (i.e. that required individuals to pay a cost when cooperating). The manner in which they evaluated or selected the individuals depended on the exact combination of team composition and level of selection. They showed that, when cooperation was needed, an homogeneous team of individuals under group selection was the best performing setting. But others have also used homogeneous teams in tasks that may require heterogeneous behaviours. For example, division of labour can be achieved between homogeneous individuals if the agents specialise during their lifetime thanks to varying initial conditions, development or environmental cues. In that context, Ferrante and colleagues (Ferrante et al. 2015) studied task partitioning, i.e. where different tasks have to be done in sequence, in the context of an evolved population of simulated foragers. They investigated the evolution of generalist behaviours (i.e. individuals who carry every task) and specialist behaviours. The teams of robots were homogeneous and they simulated evolution both with pre-evolved building blocks (i.e. pre-adapted basic behaviours) and de-novo starting only from low-level behavioural primitives. They showed that specialisation could evolve based on environmental information which was used by the agents to dynamically assign roles. Additionally, they demonstrated that particular environmental conditions, in their case a slope which made task partitioning more useful, could affect the evolution of specialists. Indeed, this slope could be used to facilitate transport and also decrease the cost of switching from one role to the other.

A particular manner in which heterogeneity is sometimes studied, especially in the field of multiagent learning, is by using Cooperative CoEvolutionary Algorithms (CCEAs) (Potter and De Jong 1994). Originally, CCEAs are used to search solutions to a given problem by decomposing this problem into subcomponents and have different populations concurrently search for a solution for each subproblem. In the context of ER, the principle is that a team is composed of robots whose controllers are separately evolved in different populations. This approach is loosely inspired by the biological coevolution of multiple species. This method is particularly useful when trying to evolve highly specialised individuals. Blumenthal & Parker (Blumenthal and Parker 2004) used a CCEA to evolve four differently abled hexapod predators in a predator-prey scenario. The issue was to ensure that the predators would evolve a correct behaviour given their own movement capabilities to prevent the prey from escaping. They used a genetic algorithm based on

punctuated learning, where the learning system is updated only after a given number of generations, and showed they could successfully evolve the predators to capture the prey. Each individual was controlled by a neural network evolved in a different population. Similarly, Yong & Miikkulainen (Yong and Miikkulainen 2009) evolved 3 agents in predator-prey scenario where the prey is faster than the predators. They used a method of neuro evolution, enforced subpopulations, which is a cooperative coevolution method where several populations of hidden neurons are separately evolved. Each of the 3 predators was encoded by neurons from a different subpopulation. The authors showed that the individuals evolved by cooperative coevolution outperformed a single centralized controller on both efficiency and robustness. Lastly, Nitschke and colleagues (Nitschke et al. 2012) introduced Collective Neuro-Evolution (CONE), a cooperative coevolutionary algorithm for the evolution of artificial neural networks, in order to evolve specialists. They evolved recurrent feed-forward networks for a team of between 50 and 100 robots, where the genotype of each robot is separately evolved. The robots had to collect blocks of different types and place them in sequence in a designated area. CONE implemented two metrics to control for specialisation and genotype similarity in order to efficiently apply recombination between populations. The goal was thus not to lose specialists during recombination. Therefore, coevolution can be used to efficiently evolve heterogeneous behaviours. However one main issue with this method is that it relies on evolving a population for each of the expected behaviours. In consequence, it needs *a priori* knowledge on the task at hand in order to structure the population accordingly. Additionally, as population size is critical in the evolutionary process, evolving several populations at the same time may be costly.

In comparison, few works have been interested in using an heterogeneous team of individuals evolved from a single population. The exact advantages of heterogeneous approaches are still not clear in the literature. Quinn (Quinn 2001) produced a comparison of homogeneous and heterogeneous approaches in a task which required two individuals to adopt coordinated motion. He compared a clonal approach, where both individuals come from the same genotype (i.e. homogeneous team), and an aclonal approach, where individuals come from different genotypes (i.e. heterogeneous team). Coordination was achieved when the individuals moved from their starting positions while staying close to one another without collision occurring. Individuals needed to take specific roles in order to coordinate efficiently. He showed that the aclonal approach outperformed the clonal approach in this task. However, Tuci & Trianni (Tuci and Trianni 2014) then ran a similar study and found different results. More precisely, they evolved a team of two robots in a setting where one of them had to stay in a designated area (the nest) and another had to move back and forth between the nest and foraging sites. As such specialisation was explicitly required to carry the task and the roles were clearly distinct. In this case, they showed that this time the clonal approach clearly outperformed the aclonal one in both efficiency and robustness. Therefore, the exact impact of genetic team composition on the evolution of cooperation is still an open issue. In particular, the manner in which specialisation is affected by homogeneous or heterogeneous teams is still debatable. The evolution of heterogeneous behaviours does not require heterogeneous control. However, it then requires that the individuals have the capabilities to dynamically specialise. In

the case of Tuci and Trianni, this was achieved thanks to a continuous time recurrent neural network. However, it may not always be desired nor cheap to endow agents with those specific capabilities depending on the context. With no *a priori* knowledge of the environment, it could be preferable to design the evolutionary process so that it may evolve efficient behaviours in simple robots. Additionally, it is not always known beforehand whether heterogeneous behaviours would even be advantageous. In conclusion, there is the need for additional works on team composition with regard to the evolution of cooperative behaviours.

5.5 Conclusions

Our goal in this Chapter was to give a brief presentation of multirobot systems and to discuss the different methods with which they could be designed. Furthermore, we wanted to highlight a critical design choice that we will explore in the two following Chapters. We showed that the design of distributed MRS, as for robotics in general, is complex and may be facilitated by resorting to automatic design. However, this is not an easy task. In particular, machine learning methods (i.e. RL) which may work for single robots do not cope well with the complexity of MRS and require to approximate the problem at hand so that it can be dealt with. In comparison, ER is another possible method which functions as a black-box optimization and thus may function well in open environment. As such we are interested in studying an open question in the field of ER: the influence of genetic team composition in the evolution of cooperative robots.

In the next two Chapters, we are interested in the evolution of cooperation among a group of heterogeneous robots. We focus on the nature of the coordination strategies evolved. Thus we consider teams of two genetically unrelated robots. While these two robots are morphologically identical, the heterogeneity in their control raises the issue of evolving cooperation when selfish behaviours can emerge. From this stems a tradeoff between using heterogeneity to evolve efficient coordination strategies and the challenge of evolving cooperation. This issue is discussed in the first Chapter. In the second Chapter, we focus solely on evolving a particular type of coordination: division of labour. Because we want this strategy to appear among heterogeneous individuals with no added capabilities, we study the issue of achieving genotypic polymorphism. Namely, we want multiple different behaviours encoded by different genotypes to coexist at the same time in a single population. In both studies, we do not use real robots. We consider our contributions here to be mainly theoretical and act as general design concepts regardless of a specific robot model. Furthermore, the transferability of evolved behaviours on real robots as well as the online evolution on physical robots are both major challenges (Floreano et al. 2008; Doncieux et al. 2015) in ER that would represent a separately substantial study for this thesis. We thus consider these issues to be beyond the scope of our work.

Chapter 6

The Tradeoff between Evolvability and Efficiency in the Evolution of Cooperation

Abstract	109
Introduction	109
Methods	110
Cooperative Foraging Task	112
Going Beyond the Evolvability vs. Efficiency Tradeoff using Incremental Evolution	115
Discussion and Conclusion	120

This Chapter is centered on the evolution of coordination behaviours in a task of collective foraging. It is organized as a published paper in an international conference:

Arthur Bernard, Jean-Baptiste André, and Nicolas Bredeche (2015). “Evolution of Cooperation in Evolutionary Robotics : the Tradeoff between Evolvability and Efficiency”. In: *Proceedings of the European Conference on Artificial Life 2015*, pp. 495–502

We are interested in the subject of evolving cooperation among a population of heterogeneous robots. In particular, we aim at comparing an homogeneous approach (i.e. a clonal approach) and an heterogeneous approach. As we previously discussed in Chapter 5, the classical approach to designing MRS in evolutionary robotics is to use a team of homogeneous robots. This is indeed the easiest way to ensure the evolution of cooperation because the fitness of the individuals is the same as that of the group. However, this means that we often focus only on the *evolvability* of the cooperative solution, i.e. the probability to evolve cooperative individuals ¹. However groups of heterogeneous robots

¹It is important to note that the term “evolvability” has a strong connotation in the field of evolutionary

may achieve greater *efficiency* (in terms of fitness performance) when coordination is required. Because individuals can adopt distinct behaviours, they could evolve more diverse coordination strategies. Yet heterogeneity may hinder the evolution of cooperation in an environment where it is possible for robots to act in a selfish manner. We thus want to compare both these approaches in such an environment on these two criteria: evolvability and efficiency.

To that end, we designed a simple collective foraging task inspired by the stag hunt. More precisely, two individuals are placed in an arena where they can forage two different types of targets: green and purple. Green targets reward the same whether they are collected alone or cooperatively. In comparison, purple targets need to be collected in a cooperative fashion but reward more than the green ones. Given this setting, we are interested in the evolution of cooperative behaviours, i.e. individuals foraging the purple targets cooperatively. We study three different approaches w.r.t. team composition:

- A control setup, where both individuals come from the same population but are genetically different.
- A clonal setup, where the two individuals are clones of each other.
- A coevolution setup, where each individual comes from a separate population.

We thus compare the results of two heterogeneous approaches (control and coevolution) and a clonal approach on the two criteria presented before.

We reveal a tradeoff between evolvability and efficiency in our foraging task. In particular, the clonal approach is shown to be the best w.r.t. evolving cooperative individuals. In comparison, the coevolution approach allows the emergence of more efficient cooperative behaviours. Furthermore, we observe in the coevolution setup the evolution of division of labour through a leader/follower strategy. We then want to overcome this tradeoff and thus improve on both evolvability and efficiency in every setup.

To that end we use incremental evolution. The individuals are first pre-evolved in a simpler cooperative task where they have to cross a set of waypoints. Rewards are obtained by crossing the same waypoints in the same order. After that, these individuals are evolved in the previous cooperative foraging task. We show that, while the probability to evolve cooperation increases in the coevolution setup, no significant differences are observed in the clonal setup. This leads to the coevolution approach attaining highest evolvability *and* efficiency compared to the other setups. However, incremental evolution implies that it is necessary to pre-evolve the individuals in the waypoints task. We thus reveal a new tradeoff: it is possible to increase the evolvability of efficient coordination strategies but at the cost of additional computational time.

robotics and may correspond to various definitions. In particular, one such definition is that it conveys the capacity for an individual to adapt to environmental change (Wagner and Altenberg 1996). This is a rising topic in artificial evolution (Banzhaf et al. 2006; Lehman and Stanley 2013; Doncieux et al. 2015; Taylor et al. 2016) and as such we do not want to create confusion with our use of the term here. In this Chapter, evolvability strictly refers to the capacity to evolve a particular trait (in our case, cooperation).

Evolution of Cooperation in Evolutionary Robotics: the Tradeoff between Evolvability and Efficiency

Abstract

In this Chapter, we investigate the benefits and drawbacks of different approaches for solving a cooperative foraging task with two robots. We compare a classical clonal approach with an additional approach which favors the evolution of heterogeneous behaviors according to two defining criteria: the evolvability of the cooperative solution and the efficiency of the coordination behaviors evolved. Our results reveal a tradeoff between evolvability and efficiency: the clonal approach evolves cooperation with a higher probability than a non-clonal approach, but heterogeneous behaviors evolved with the non-clonal approach systematically show better fitness scores. We then propose to overcome this tradeoff and improve on both of these criteria for each approach. To this end, we investigate the use of incremental evolution to transfer coordination behaviors evolved in a simpler task. We show that this leads to a significant increase in evolvability for the non-clonal approach, while the clonal approach does not benefit from any gain in terms of efficiency.

Introduction

The evolution of cooperative actions in evolutionary robotics is as much a challenge as an interesting perspective for the design of complex collective systems (Doncieux et al. 2015). As such, it has been widely studied with very diverse approaches and objectives (Waibel et al. 2009; Hauert et al. 2014; Trianni et al. 2007; Lichocki et al. 2013). These works often use a clonal paradigm, where each robot has a copy of the same genome. This makes sense as this is the easiest way to ensure cooperation when individuals are expected to display similar behaviors. Moreover, using clones ensures minimal genetic relatedness between individuals, which is known to allow the evolution of altruism (Waibel et al. 2011; Montanier and Bredeche 2011). As such, most research focus on increasing the probability for the cooperative solution to evolve.

In comparison, the nature of coordination behaviors and their influence on the quality of cooperation has yet to be thoroughly studied. In particular, interactions between clones in evolutionary robotics tend to produce homogeneous behaviors when most coordination tasks could benefit from heterogeneous behaviors. This could be solved by using a non-clonal approach where paired individuals do not use the same genome, and could possibly evolve different behaviors more easily. However, a non-clonal approach may face a *chicken-and-egg* dilemma: multiple individuals need to behave in a particular fashion for cooperation to be rewarding, but no benefit can be extracted from this behavior unless all individuals cooperate. Therefore, without cooperating partners, those behaviors cannot be selected by the evolution as they do not benefit the individual. This

is particularly problematic when a moderately rewarding solitary strategy overshadows a more rewarding, but also more challenging to evolve, cooperative strategy (Skyrms 2004).

In this Chapter, we are interested in the comparison between clonal and non-clonal approaches on two different criteria:

- *Evolvability*² of cooperation, which is the number of successful runs where cooperation evolved.
- *Efficiency* of cooperation. This criteria is focused on the quality of the evolved behaviors and is determined by the performance (w.r.t. fitness score) of the coordination strategies.

To that end, we design a foraging task where both cooperative and solitary strategies are possible but where cooperation provides the largest reward. This task is favored by the evolution of efficient cooperative behaviors and we compare different approaches on both criteria. The first approach is a straightforward implementation of the literature where interacting individuals are clones. In comparison, the second approach is a rather extreme implementation of a non-clonal approach: we use coevolution, where individuals are from two different populations, and where fitness scores are computed independently for each individual. While this scheme is typical of competitive coevolution (Floreano and Nolfi 1997; Floreano et al. 1998; Panait and Luke 2005), the nature of the task considered here makes cooperation more interesting, as both individuals can selfishly benefit from being cooperative.

In the next section, we describe the methods and experimental setup used throughout our study. Then, we compare the results of the two approaches on the cooperative task. This first experiment reveals that both approaches face a tradeoff between evolvability and efficiency, where neither one dominates the other on both criteria. We investigate in a second experiment the possibility to overcome this tradeoff for both approaches. To this end, we use incremental evolution (Harvey et al. 1994; Urzelai et al. 1998) and evolve coordination in a simpler task in order to improve both the evolvability and efficiency on the target task for each approach. Finally, we discuss the implication of our findings in the last section, in particular with respect to maximizing evolvability and efficiency alike.

Methods

Two robotic agents are placed in a 800 by 800 units square arena with four solid walls and emptied from any obstacle apart from the targets in the foraging task. Each circular-shaped agent, with a diameter of 20 units, has a collection of sensors divided between a 90 degrees front camera and 12 uniformly distributed proximity sensors. The camera is composed of 12 rays with infinite range which indicate the type (coded on 3 bits) and proximity (one value in R^n) of the nearest object or agent in their direction. Proximity sensors have a range of twice the agent body's diameter and are used to get the distance

²As a reminder, we restrict here the definition of evolvability to the capacity to evolve cooperative solutions.

to any obstacle nearby such as solid objects, the other agent or walls. The two agents always begin the simulation next to one another at one end of the arena, whereas all the objects' initial positions are randomized.

Agents can move freely in the environment and are controlled by a fully connected multi-layer perceptron with a single hidden layer, the topology of which does not change during the evolution. Inputs of this neural network are fed with all the data extracted from the sensors: 48 neurons for the camera (4 neurons for each of the 12 rays) and 12 neurons for the proximity sensors. A bias neuron, whose value is always 1, brings the total number of input neurons to 61. The hidden layer is comprised of 8 neurons and the output layer of 2 neurons giving the speed of each of the agent's wheels. The activation function used is a sigmoid.

In each experiment, individuals evolved during a fixed amount of evaluations thanks to an evolutionary algorithm. Their genome consists of a collection of the 506 connection weights (as real-values) of the neural network and is initially randomized for each individual in the population. Three evaluation setups are used to compare the different approaches of our experiment:

- In the *control* setup, each individual is evaluated against 5 other randomly chosen individuals in the population except itself. Therefore we ensure that there is no genetic relatedness between individuals in each pair. However, it is not clear how the evolutionary algorithm itself may impact the population's diversity, especially because elitism is used;
- In the *clonal* setup, each individual is evaluated once against a clone of itself. This setup is used to study the results of the classical clonal approach (Waibel et al. 2009; Hauert et al. 2014; Trianni et al. 2007; Lichocki et al. 2013). While previous works have shown on multiple occasions that cooperation can evolve, it is not clear if individuals can take different roles during a cooperative interaction;
- In the *coevolution* setup, each of the two individuals comes from two different coevolved populations. In this setup, each individual from one population encounters 5 random individuals from the other population. As pairing considers individuals from two separate populations, the evolution of heterogeneous behaviors is theoretically easier. As a matter of fact, such a relation where two very different individuals find a selfish interest in mutual cooperation is actually quite common in nature (Connor 1995).

A pair of individuals then interact in the arena described before for a fixed number of simulation steps called a trial. Each trial is conducted 5 times to account for the random initial positions of the objects and decrease the influence of the initial conditions on the individuals' performance.

The selection method used in the evolutionary algorithm is an elitist (10+10)-ES where the 10 best individuals in the population are used to generate 10 offsprings for the next generation. We use no recombination and therefore each offspring is a mutated copy of its parent. Mutations were sampled according to a gaussian operator

with a standard deviation of 1.10^{-2} and a gene's mutation rate of 5.10^{-3} . Finally, population size was kept constant through the evolution with a number of 20 individuals. All experiments were done using the framework for evolutionary computation SFERESv2 (Mouret and Doncieux 2010), which includes a fast 2D robotic simulator. The source code for reproducing the experiments is available for download at <http://pages.isir.upmc.fr/~bredeche/Experiments/ECAL2015-coop.tgz>.

Cooperative Foraging Task

In this first experiment, we investigate the evolution of cooperation in a foraging task. The environment is filled with 18 solid targets that the agents can collect. To collect a target, an agent has to stay close to this object for a fixed amount of simulation steps (800). After this duration, the target disappears and any agent close to it is rewarded with its value. Targets are of two types. Green targets always reward 50 when collected whereas purple ones reward 250 only when the agents collect it together (Table 6.1). If a solitary agent collects a purple target, it disappears and rewards nothing. Consequently, there is both an incentive and a risk to cooperate as cooperation is dependent on successful coordination. This setup is a robotic implementation of a well-known problem in game theory for studying the evolution of mutualistic cooperation: the *Stag Hunt* (Skyrms 2004).

The fitness score (F) of an individual is the average reward per trial:

$$F = \frac{1}{N * M} \sum_{i=1}^N \sum_{j=1}^M f_{ij}$$

Where N is the number of individuals encountered (5 in the control and coevolution setups, 1 in the clonal setup), M the number of trials (5) and f_{ij} the rewards obtained at trial j with individual i .

When a target is collected, another target of the same type is then placed at a random position in the arena to keep a constant ratio between green and purple targets. Each evaluation lasted 20000 simulation steps and 60 independent runs were conducted for each experimental setup, each one lasting 40000 evaluations.

Target		Reward
Green		
	<i>alone</i>	50
	<i>coop</i>	50
Purple		
	<i>alone</i>	0
	<i>coop</i>	250

Table 6.1 – *Rewards for the foraging of the different targets. Rewards depend on whether they were collected alone or cooperatively.*

Results

Setting	# Coop.	# Solitary	Total
Control	10	50	60
Clonal	28	32	60
Coevolution	14	46	60

Table 6.2 – *Evolution of a cooperative strategy.* Number of simulations where the best individual evolved a cooperative strategy (collecting purple targets) or a solitary strategy (collecting green targets) for each setup in the foraging task.

We are interested in the number of simulations where cooperation evolved (i.e. the *evolvability* of each approach), which means simulations where the best individual in the population evolved the cooperative foraging of the purple targets (i.e. more than 50% of the collected targets are purple). Results for the three setups are displayed in Table 6.2. As could be expected from the literature, the clonal setup displays a greater evolvability w.r.t. evolving cooperation (28/60), whereas coevolution (14/60) is on par with the control setup (10/60). It is also apparent that cooperation is still difficult to evolve as in the best case (clonal), no more than half the simulations display the evolution of cooperative behaviors.

However, cooperative individuals do not perform with the same *efficiency* from one setup to another. We show in Figure 6.1 the median fitness score of the best individuals in each independent run where cooperation evolved over time and for each setup. Fitness scores are significantly different in each setup with the best score obtained in the coevolution setup and the worst in the control setup (Mann-Whitney U-test on the fitness score of the best individuals at the last evaluation, p -value < 0.001).

These differences in efficiency can be explained by looking at the nature of the cooperative behaviors evolved, which reveals two types of behaviors: *turning* and *leader/follower*.

Individuals adopting the turning strategy turn around one another so that they always see the other individual as well as stay close to it (Figure 6.2(A)). This allows the two individuals to approach simultaneously a same target and therefore forage it in a cooperative fashion. In this strategy, both individuals have a similar behavior and no role division is necessary for their successful cooperation.

In comparison, individuals which evolve a leader/follower strategy adopt a differentiation between two roles: *leader* and *follower* (Figure 6.2(B)). The individual we call leader always goes first on a target whereas the follower always arrives second on the same target. We observe that the follower’s behavior consists in staying close to the leader and always keeping it in front of itself. In comparison the leader shows a lesser interest in the presence of its follower and rarely checks on its position.

Table 6.3 shows the distribution of cooperative strategies for all three setups. Whereas the control and clonal setups always resulted in turning strategies (resp. 10/10 and 28/28), the coevolution setup always displayed the evolution of a leader/follower strategy (14). We observe that this latter strategy leads to more efficient cooperation. Indeed, individuals adopting the turning strategy are forced to check constantly on the other individual’s

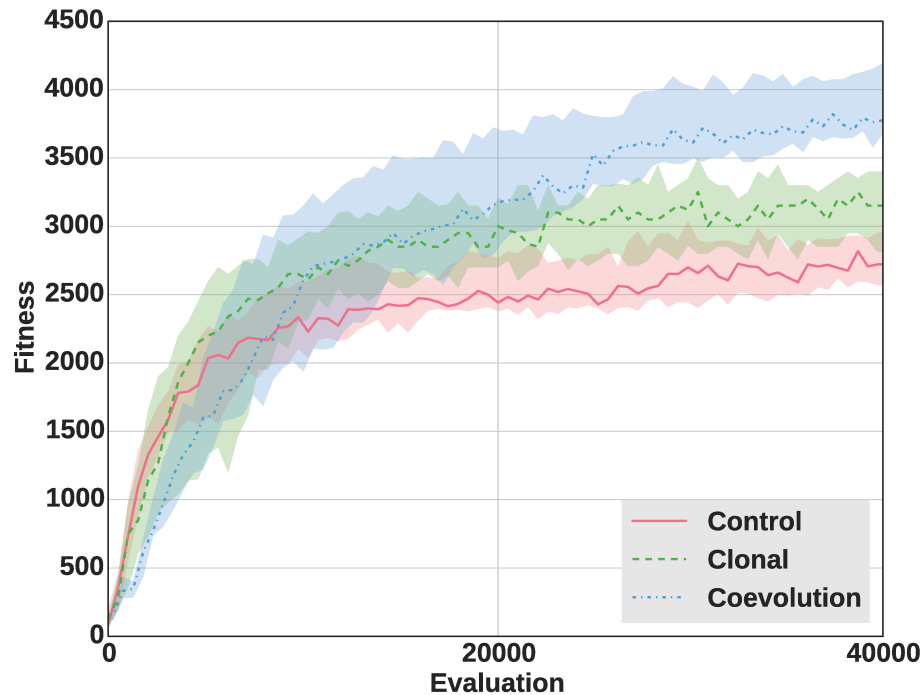


Figure 6.1 – Performance of the cooperative solutions. Median fitness score of the best individuals in each of the runs where cooperation evolved for each setup over time. The fitness score of an individual is computed as the average reward the individual earned per trial by foraging targets. The colored areas around the medians represent the first and third quartiles.

position. Consequently, they cannot be as fast as individuals with a leader/follower strategy where they move to the target in a straight line under the leader’s direction. Moreover, due to the random proximity of the targets, the turning strategy is prone to errors. Namely, they often get to another target than that of their partner whenever two targets are too close to each other.

A possible explanation as to why no leader/follower strategy could evolve in the control and clonal setups may be because of the need to differentiate between the two roles. Indeed, there needs to be the existence of an asymmetry between the two individuals for this phenomenon to appear. With coevolved populations, this asymmetry is deliberately created by the separation between the two populations. Indeed, we observe that one population exclusively contains leaders while the other exclusively contains followers.

The two other setups fail to evolve heterogeneous behaviors. In the control setup, this may be due to the evolutionary algorithm used, especially with elitism enforcing the homogenization of the population throughout the course of evolution (as hinted in the Methods Section). Then, the clonal setup introduces yet another challenge as switching to a particular role can only be done during evaluation as both individuals are by definition genetically similar.

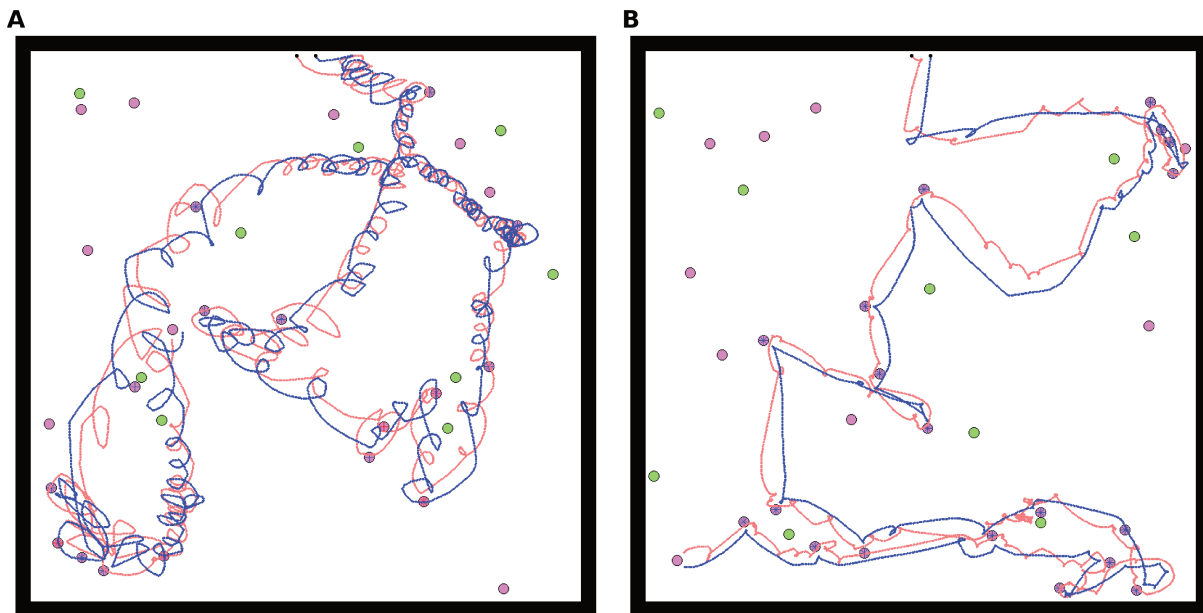


Figure 6.2 – *Snapshots of the simulation after an entire trial in the foraging task. The path of each robotic agent from their initial positions (black dots) is represented in red and blue. The green and purple discs represent the 18 targets in the environment. When a target is foraged by the two agents, a red cross (resp. blue) is drawn on the target if the red agent (resp. blue) arrived on it first. Each snapshot corresponds to a trial where agents adopted a different behavior: (A) turning or (B) leader/follower.*

Going Beyond the Evolvability vs. Efficiency Tradeoff using Incremental Evolution

The previous section revealed a tradeoff between evolvability and efficiency. In the clonal setup, cooperation evolves more often than with other setups. However, the coevolution setup yields cooperative behaviors which are more efficient, with paired individuals displaying asymmetrical behaviors.

In this section, we address the following question: is it possible to benefit from both evolvability *and* efficiency with the clonal and/or the coevolution setups? In other words, we explore (1) whether the clonal setup can be used to evolve pairs with heterogeneous behaviors, and (2) whether the coevolution setup can be improved in terms of number of runs where cooperation evolved.

In order to address this question, we use incremental evolution, a rather common method in evolutionary robotics for solving challenging problems (Dorigo and Colombetti 1994; Saksida et al. 1997; Bongard 2008; Doncieux 2013). The main principle is to ease the learning of a complex task by splitting it into simpler sub-tasks (Perkins and Hayes 1996).

In the following, we introduce an additional task, the *waypoints crossing* task, which requires the evolution of coordination behaviors, and is simpler to address than the previous task. Individuals evolved in this first task are then used as starting point for the original

Setting	# Leader/Follower Strategy	# Turning Strategy	Total
Control	0	10	10
Clonal	0	28	28
Coevolution	14	0	14

Table 6.3 – Evolution of a cooperative strategy. Repartition of the different strategies evolved in each of the runs where cooperation evolved for each setup in the foraging task. We indicate in each cell the number of simulations where a particular strategy evolved.

task described earlier, hoping that cooperative behavior will be *recycled* from the first task to the second task.

Waypoints Crossing Task

We consider a task where robotic agents have to cross randomly positioned waypoints. As such, these round waypoints do not act as obstacles and have a diameter of 30 units. As soon as an agent goes through a waypoint, it can not be seen by this agent anymore. All 18 waypoints have the same color and can be crossed in any order. The fitness score (F) of each individual is defined as the average longest sequence of waypoints shared by both agents per trial:

$$F = \frac{1}{N * M} \sum_{i=1}^N \sum_{j=1}^M l_{max_{ij}}$$

Where N is the number of individuals encountered (5 in the control and coevolution setups, 1 in the clonal setup), M the number of trials (5) and $l_{max_{ij}}$ the longest sequence of waypoints shared by both individuals at trial j with individual i .

This implies that the two individuals are rewarded when crossing waypoints in the same order as well as maximizing the number of waypoints crossed. Each evaluation lasted 10000 simulation steps and 60 independent runs were conducted for each experimental setup, each one lasting 40000 evaluations.

All simulations showed an increase in fitness score for each of the three setups (cf. Figure 6.3). This was expected as this task does not represent a particular challenge for the individuals: it simply needs the evolution of a successful coordination strategy. However, whereas the coevolution and clonal setups performed equally, they both surpassed the performance of individuals from the control setup (Mann-Whitney, p -value < 0.001).

As with the previous foraging task, we can hypothesize that these differences in fitness scores are due to differences in the behaviors evolved. Table 6.4 gives a classification of the cooperative behaviors for each setup. They are similar to those in the previous task with the addition of a third rare strategy: the *wall-following* strategy (which is regrouped in “Other”). Wall-followers simply follow the walls around the arena and cross any waypoints close to the wall they are adjacent to. As such, this is a far less efficient strategy than the two others.

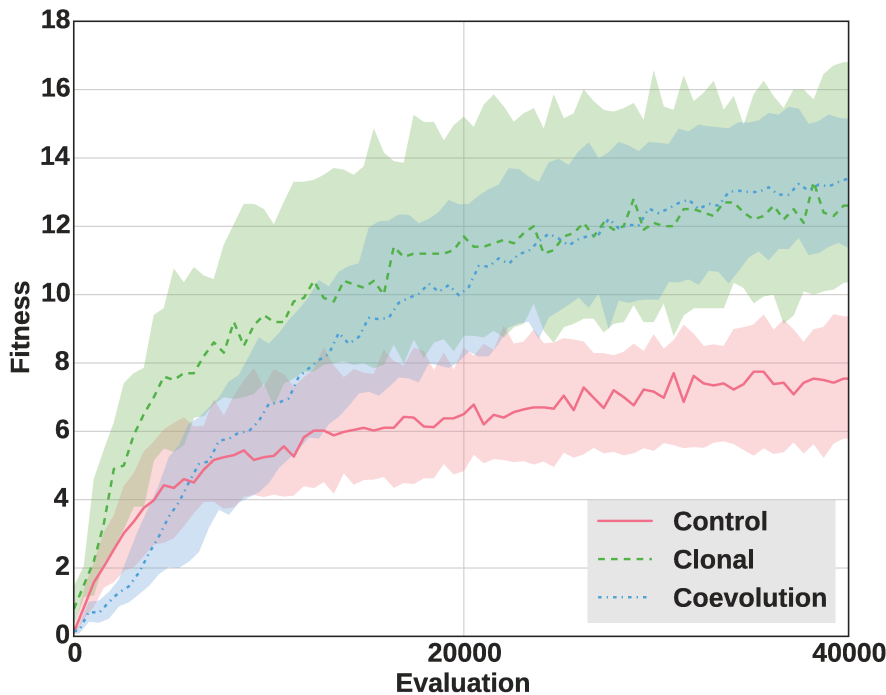


Figure 6.3 – Performance of the cooperative solutions. Median fitness score of the best individuals in each of the 60 independent runs and for each setup over time. Fitness score is computed as the average longest sequence of waypoints shared by both agents per trial. The colored areas around the medians represent the first and third quartiles.

In the coevolution setup, nearly all runs (59/60) evolved a leader/follower strategy. Interestingly, although fitness scores in the clonal and control setups are significantly different, this behavior evolved in roughly one third of the runs for both setups. To explain the difference in fitness scores, we must take into account the quality of the leader/follower strategy in each setup. We measure the proportion of leadership in each interaction, which is computed as the proportion of waypoints crossed by both individuals for which the leader arrived first. Figure 6.4 displays the boxplots of the proportion of leadership for the best individuals in each setup and only for the simulations where a successful leader/follower strategy evolved (a minimal threshold of 0.75 is chosen to consider only the best performing runs). We show that the proportion of leadership is greater in the clonal and coevolution setups than in the control one (Mann-Whitney, p -value < 0.005). These differences means that the individuals are more efficient in their leader/follower strategy in the clonal and coevolution settings than in the control one. This explains the differences in fitness scores observed in Figure 6.3.

Interestingly, whereas in the foraging task no leader/follower strategy could evolve in the control and clonal setups, this strategy did evolve in one third of the simulations for this task. This could mean that these individuals use information in the environment to adopt one role or the other. Indeed, we observe that this is achieved by exploiting the differences in the initial starting positions, with one individual on the left and the other

Setting	# Lead.	# Turn.	# Other	Total
Control	19	37	4	60
Clonal	23	31	6	60
Coevolution	59	1	0	60

Table 6.4 – Cooperative strategies evolved. *Repartition of the different strategies evolved in each of the 60 independent runs for each setup in the waypoints task. We indicate in each cell the number of simulations where a particular strategy evolved: Leader/follower (Lead.), Turning (Turn.) or Other. “Other” regroups wall-following strategies or simulations where no recognizable strategy evolved.*

on the right. They both turn to the same direction (left or right, depending on the runs) at the beginning of the simulation which results in one individual (the leader) turning its back to the other, while the second individual (the follower) looking at its partner.

Recycling Cooperative Behaviors in the Foraging Task

Coming back to the initial foraging task, we perform the exact same experiment described at the beginning of this Chapter, with one notable exception: the initial population is initialized with genomes evolved for solving the waypoint task. This implies that coordination is possible starting from the very first generation of each setup. Given that we have already shown that such coordination is a desirable feature, the question is: will it be possible to retain cooperative behaviors in order to solve the foraging task?

Setting	# Coop.		# Solitary	Total
	# Lead.	# Turn.		
Control	0	20	40	60
Clonal	0	24	36	60
Coevolution	28	0	32	60

Table 6.5 – Evolution of a cooperative strategy. *Proportion of the 60 independent simulations where the best individual evolved a cooperative strategy (collecting purple targets) or a solitary strategy (collecting green targets) for each setup in the foraging task when individuals are previously evolved in the waypoints task. In addition, the repartition of the different strategies is indicated when cooperation evolved: Leader/Follower (Lead.) or Turning (Turn.).*

Table 6.5 gives the results in terms of evolved behaviors from the 60 independent runs for each setup. The coevolution setup evolves cooperation slightly more often (28/60) than both the control (20/60) and the clonal (24/60) setups. A first remark is that the number of occurrences of cooperation for the coevolution and control setups have actually doubled compared to previous results without incremental evolution (see Table 6.2). This is not the case for the clonal setup, which does not appear to benefit from incremental evolution.

A second remark is that cooperation in the coevolution setup systematically corresponds to a leader/follower strategy, which is *never* the case with the two other setups. This has a

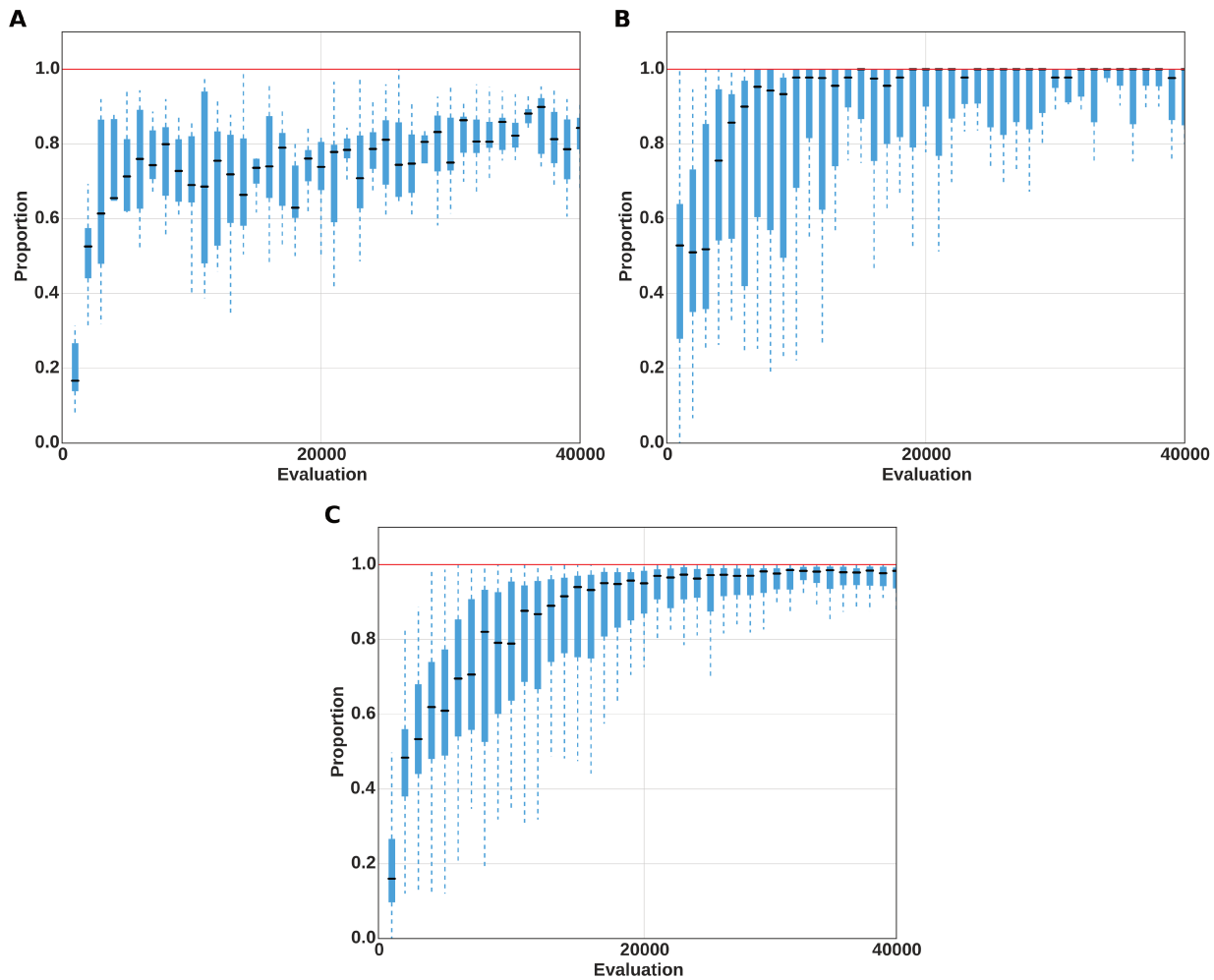


Figure 6.4 – Proportion of leadership. Boxplots of the proportion of leadership over time for the best individuals in each runs where the proportion at the last evaluation was greater than 0.75 in the (A) control, (B) clonal or (C) coevolution setup. This value represents the proportion of waypoints crossed by both individuals for which the leader arrived first.

significant, though expected, impact on fitness scores, as shown in Figure 6.5. Cooperation evolved with the coevolution setup leads to significantly greater fitness scores (Mann-Whitney, p -value < 0.001).

Results from this experiment make it possible to revise our initial statement. Using pre-trained individuals strongly benefits the coevolution setup in terms of evolvability. But this is not the case with the clonal setup, for which using pre-trained individuals improves neither evolvability nor efficiency. Therefore, we may face a tradeoff which does not concern evolvability and efficiency, but one that implies computational cost: the coevolution setup outperforms the clonal setup on both evolvability and efficiency *at the cost of additional computational effort*.

The control and clonal setups completely failed to maintain a leader/follower strategy, even though such strategy originally evolved. An explanation is provided by considering

the difference between the waypoints task, where leader/follower evolved, and the current foraging task. In the waypoints task, symmetry breaking could be achieved at the beginning of the evaluation (as explained earlier), and could be retained afterwards as the follower was always behind the leader. However, the current foraging setup requires that the two robots display the same behavior to cooperatively collect a target (ie. both robots have to touch the target), which implies that leader/follower roles are lost, as they depend on the relative position of robots with one another.

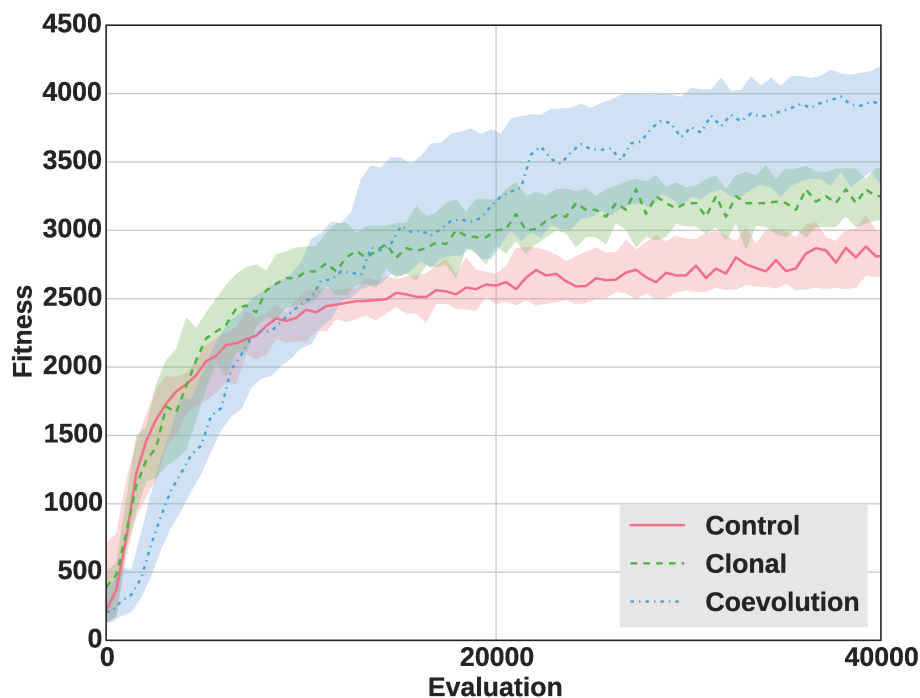


Figure 6.5 – Performance of the cooperative solutions. Median fitness score of the best individuals in each of the runs where cooperation evolved for each setup over time. The fitness score of an individual is computed as the average reward the individual earned per trial by foraging targets. The colored areas around the medians represent the first and third quartiles.

Discussion and Conclusion

In this Chapter, we considered several approaches for the evolution of cooperation in evolutionary robotics: a clonal approach, where all individuals in a group share the same genome, and a non-clonal approach, where individuals are independent from one another, but may share a common interest in cooperating.

We first showed that there exists a tradeoff between evolvability and efficiency. On the one hand, the clonal approach evolves cooperative behaviors on a more frequent basis than with the other approach. On the other hand, the non-clonal approach, which is implemented using a coevolution setup, results in more efficient behaviors in terms

of pure performance whenever cooperation evolved. The non-clonal approach actually enables the evolution of asymmetric behaviors, such as a leader/follower strategy.

We then used incremental evolution to evolve coordination behaviors using a simpler task in order to overcome this tradeoff and improve both evolvability and efficiency in each setup. We showed that while no improvement was observed in the clonal setup on either criteria, the outcome is very different for the coevolution setup: the probability of evolving cooperation actually increases, and the evolved cooperative solutions remain the most efficient.

This work raises several questions. Firstly, heterogeneous behaviors were obtained with coevolution, a rather radical way to enable asymmetrical behaviors during cooperation. However, the waypoints task revealed that breaking symmetry can also be done with identical individuals using environmental feedback, even though such cooperation is difficult to obtain. As a consequence, the evolution of cooperation with heterogeneous behavior without resorting to coevolution could be investigated. In particular, we could study how more elaborated neural architectures (e.g. using plasticity) can switch to a particular persistent regime depending on environmental cues available at the beginning of the evaluation.

Secondly, incremental evolution requires an added computational cost in order to increase evolvability in the non-clonal approach. However, it may be possible to avoid this extra cost by considering other evolutionary methods. In particular, a multiobjective approach which considers both performance and *diversity* could improve the optimization process (Lehman and Stanley 2008; Doncieux and Mouret 2014). Though this approach looks promising, it is not clear yet how diversity should be implemented in the context of cooperative problem solving.

Chapter 7

The Evolution of Specialisation through Genotypic Polymorphism

Abstract	125
Introduction	125
Methods	126
Behaviours of Specialists in a Cooperative Foraging Task	128
Evolving Heterogeneous Behaviours with an Elitist Selection	130
Evolution Under a Fitness-Proportionate Selection	132
Computational Analyses of Population Dynamics	134
Key Properties for Evolving Heterogeneous Behaviours	136
Discussion and Conclusions	137

We now investigate the evolution of division of labour between heterogeneous robots via genotypic polymorphism. Results are presented as a published article in an international conference:

Arthur Bernard, Jean-Baptiste André, and Nicolas Bredeche (2016a). “Evolving Specialisation in a Population of Heterogeneous Robots: the Challenge of Bootstrapping and Maintaining Genotypic Polymorphism”. In: *Proceedings of the Artificial Life Conference 2016 (ALIFE XV)*, pp. 152–159

In the previous Chapter we showed that it could be beneficial to consider the quality of cooperative behaviours in addition to the probability to evolve cooperation. In particular, we revealed that a particular heterogeneous approach, cooperative coevolution, led to the emergence of a more efficient coordination behaviour: leader/follower. This behaviour entails the evolution of task specialisation (or division of labour). Here we focus on the evolution of a leader/follower strategy in a single population of individuals.

Because in evolutionary robotics cooperation is often evolved among homogeneous robots, this means that in order to achieve division of labour, individuals must be capable to dynamically allocate their roles during their lifetime. In consequence, specialisation often relies on neuronal plasticity or environmental cues. Here we are interested in



evolving specialisation without those additional mechanisms. This means that we want to achieve division of labour at the level of the population. In consequence, we want to investigate the maintenance of several different genotypes encoding for different roles in a single population, i.e. genotypic polymorphism. In particular, we focus on the impact of selection strategies on evolving genotypic polymorphism.

To that end, we use a simpler foraging task than the one presented in Chapter 6. Two genetically different individuals are placed in arena where they can collect a single type of targets. This target is more rewarding when collected in a cooperative manner. This means that evolving cooperation is easy (we are not interested in the issue of evolving cooperation here) and that the task is favored by the evolution of efficient coordination strategies. As we showed in the previous Chapter, two different cooperative strategies can evolve in this setting. On the one hand, there is the turner strategy, where both individuals adopt the same behaviour to coordinate. Therefore this is a generalist behaviour. On the other hand, they can also evolve specialist behaviours and adopt a leader/follower strategy. We are interested on studying the evolutionary differences of two selection schemes: (1) a $(\mu + \lambda)$ -ES (elitist) selection strategy and (2) fitness-proportionate selection. We also study the impact of varying population sizes.

We reveal that specialisation is nearly impossible to evolve under an elitist selection strategy. Indeed, in only one replication under high population size do we observe the presence of specialists in the population at the end of evolution. Surprisingly however, specialists do appear in multiple replications but are never maintained. Even when the population is initially seeded with specialists, similar results are observed. In comparison, while specialists nearly never evolved under fitness-proportionate selection, they are easily maintained throughout evolution (especially when starting with a population of specialists). We thus reveal that evolving genotypic polymorphism is hindered by the challenge of both evolving and maintaining specialists in the population and that none of these two classical selection schemes are suitable to that end.

We then use computational analyses to garner a deeper understanding of the underlying dynamics at play. We show that while specialists are the most efficient, generalists can invade the population because they fare quite well against any other phenotype. In comparison, specialists need to be paired with other specialists of a different type to perform adequately. Additional analyses show that, under finite population size genetic diversity can be lost from one generation to the other. This can, especially with small population sizes, lead to the disappearance of specialists from the population. We thus highlight two critical properties for the evolution of genotypic polymorphism: (1) protection against the invasion of generalists and (2) maintenance of genotypic diversity. We argue that an algorithm endowed with such properties would enable genotypic polymorphism to be achieved.

Evolving Specialisation in a Population of Heterogeneous Robots: the Challenge of Bootstrapping and Maintaining Genotypic Polymorphism

Abstract

In this Chapter, we are interested in the evolution of specialisation among a single population of heterogeneous robotic agents in a cooperative foraging task. In particular, we want to compare (1) the emergence and (2) fixation of genotypic polymorphism under two different selection methods: elitist and fitness-proportionate. We show that, while the emergence of specialists is easy under an elitist selection, this method cannot maintain heterogeneous behaviours throughout the whole simulation. In comparison a fitness-proportionate algorithm proves to be inefficient in evolving any cooperative strategy but ensures the conservation of heterogeneity when it is present in the population. We then reveal through additional experiments two key factors for the evolution of heterogeneous behaviours in our task: (1) protection of genotypic diversity and (2) efficient selection of partners. We finally demonstrate this assertion and, while our main problem remains unsolved, we provide directions on how it could be successfully approached.

Introduction

Task specialisation is a defining characteristic in achieving efficient coordination and is thus considered to be crucial in the evolution of complex cooperative behaviours (Szathmàry and Maynard Smith 1995). The problem of evolving cooperation has been largely studied in evolutionary robotics as it raises interesting perspectives for the design of collective robotics (Trianni et al. 2007; Hauert et al. 2014; Doncieux et al. 2015). As a consequence, the manner in which robotic agents could evolve specialisation (or division of labour) for a cooperative task represents a compelling challenge in evolutionary robotics. As such, a large body of literature has already been dedicated to this subject. However, most research focus on the particular case of homogeneous groups of individuals (Waibel et al. 2009) as is classic in evolutionary robotics. This means that the individuals are forced to rely on phenotypical plasticity (Waibel et al. 2006; Ferrante et al. 2015; Eskridge et al. 2015) and/or environmental cues (Waibel et al. 2006; Goldsby et al. 2010) in order to achieve specialisation.

In this Chapter, we focus on a slightly different problem: the evolution of a polymorphic population where division of labour is encoded at the genotypic level. More precisely, we want to study the evolution of a population containing two (or more) different types of genotypes. Each of these types of genotype should be able to encode for a different

role without requiring the addition of mechanisms for lifetime specialisation. Thus it poses the problem of both *evolving* and *maintaining* genotypic polymorphism in a single population. Here we want to investigate the conditions under which specialised behaviours for a cooperative task can evolve in a single population of heterogeneous individuals. In particular, we are interested in the influence of the selection process in achieving division of labour.

We design a 2-robots cooperative foraging task where both a solitary and a cooperative strategies can evolve but where cooperation is highly rewarded. The genotype of each robotic agent is separately chosen in the population and the individuals therefore form an heterogeneous group. This task is greatly favored by the evolution of efficient coordination strategies. In particular, our previous work on a similar task (Bernard et al. 2015) showed that two types of cooperative strategy could evolve: one where both individuals adopt homogeneous behaviours (generalists) and the other one where they adopt a leader/follower strategy (specialists). Moreover, it was shown that the latter could only emerge between heterogeneous individuals. As it is also the more efficient behaviour, we study the conditions for its emergence. The evolutionary dynamics of two popular selection methods are studied: (1) an elitist ($\mu + \lambda$) evolution strategy and (2) fitness-proportionate selection. Fitness-proportionate in particular is interesting with regards to genotypic polymorphism as it is known to allow the evolution of frequency-dependent selection (Altenberg 1991).

In the next Section, we introduce the experimental setup. Then we present the two types of cooperative strategies that can evolve. Next, we investigate whether any of the selection methods could evolve heterogeneous behaviours. In particular, we study for both schemes the evolutionary outcomes depending on whether the population is initially constituted of random individuals or seeded with pre-evolved efficient specialists. Then we present the results of computational analyses in order to reveal and understand more deeply the mechanisms at play. In a final experiment, we reveal key mechanisms which could be investigated to solve this problem. Finally we discuss our findings and shed light on interesting perspectives for future work.

Methods

We evaluate two robotic agents in a 800 by 800 units square arena devoid of any obstacles except for the foraging targets. At the beginning of a simulation, 18 targets are randomly positioned in the environment. While the agents may move freely in the arena, the targets' positions are fixed. For a target to be collected, any agent needs to stay in contact with it for a specified amount of time (800 simulation steps). The target is removed after this duration and put back at another random position so that the number of targets is kept the same throughout a simulation. We consider that cooperative foraging happens if both individuals are in contact of the target when it is removed. When an agent collects a target, it is rewarded **50** if this target has been foraged in a solitary manner or **250** if both agents have cooperated to collect it.

Each agent is circular-shaped with a diameter of 20 units and possesses a collection of different sensory inputs. The first type of inputs is a 90 degrees front camera and

Foraging	Reward
Solitary	50
Cooperatively	250

Table 7.1 – Rewards for the foraging of targets. Rewards depend on whether the targets were collected in a solitary or cooperative fashion.

is composed of 12 rays, each one indicating the type and distance to the nearest object (either another agent or a target). The other type of inputs are 12 proximity sensors evenly distributed around the agent’s body. With a range of twice the agent’s diameter, each proximity sensor outputs the proximity of the nearest obstacle in its range.

Both agents begin the simulation next to each other at the same end of the arena and can move according to the outputs of their neural network. This neural network is a fully connected multi-layer perceptron with one hidden layer. The inputs of the neural network are comprised of all the sensory information of the agent, i.e. 36 input neurons for the camera (3 inputs for each ray) and 12 for the proximity sensors. A final input neuron whose value is always 1 is used as a bias neuron. This amounts the total number of input neurons to 49. The hidden layer is constituted of 8 neurons while the 2 neurons of the output layer return the speed of the agent’s wheels. A sigmoid is used as the activation function of each neuron. Finally, the topology of the network is kept constant during the experiments.

The population of individuals is evolved thanks to a classical evolutionary algorithm. The genotype of each individual is constituted of a collection of the 410 real-valued connection weights of the neural network. At each generation of the algorithm, every individual is evaluated by being successively paired with another individual randomly chosen in the population 5 times. Each pair interacts in the setting presented before during 20000 simulation steps which we call a *trial*. We perform 5 trials for each pair of individuals in order to decrease the impact of the targets’ random positions on the individuals’ performance. The fitness score of an individual is computed as the average reward per trial.

The population for the next generation is created according to two different selection schemes :

- **$(\mu + \lambda)$ elitist selection:** the population of the next generation is constituted of the μ best individuals from this generation and λ offsprings sampled from the best individuals.
- **Fitness-proportionate:** offsprings are randomly sampled from the current generation to constitute the population of the next generation. The probability to sample a particular parent is proportional to this parent’s fitness score.

Regardless of the selection method used, every offspring is a mutated clone of its parent and no recombination is used in our algorithm. The probability for each gene to mutate is 5×10^{-3} and mutations are sampled according to a gaussian operator with a standard deviation of 2×10^{-2} . Finally, experiments were conducted with the robotic

2D simulator of SFERESv2 (Mouret and Doncieux 2010), a framework for evolutionary computation. You can find the source code for the experiments available for download at <http://pages.isir.upmc.fr/~bredeche/Experiments/ALIFE2016-specialisation.tgz>.

Behaviours of Specialists in a Cooperative Foraging Task

We showed in a previous article (Bernard et al. 2015) that two cooperative strategies could evolve in this particular task: *turning* (between two *turners*) and *leader/follower* (between a *leader* and a *follower*). Both of these strategies achieve cooperative foraging but with varied efficiency.

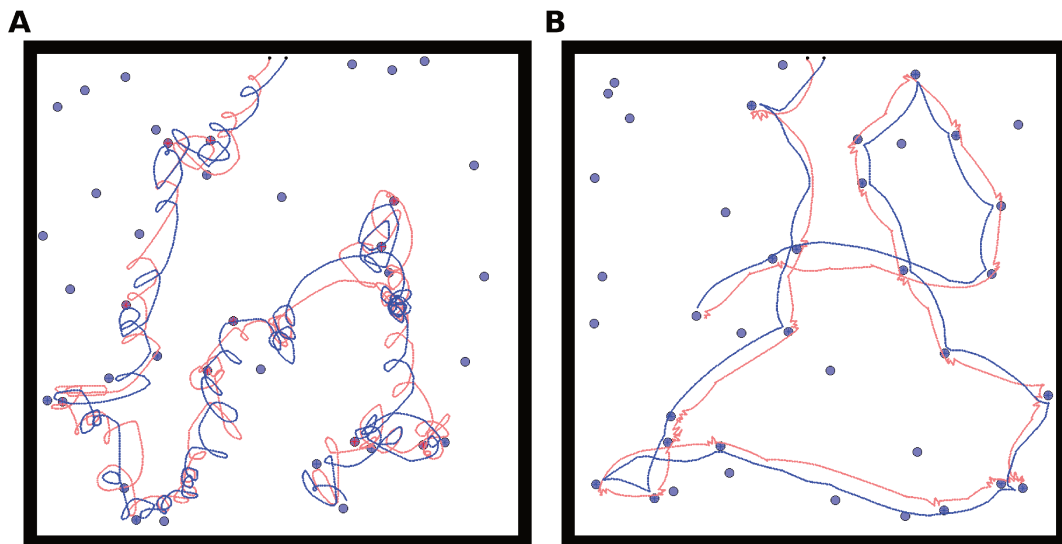


Figure 7.1 – Snapshots of the simulation after an entire trial in the foraging task. The path of each robotic agent from their initial positions (black dots) is represented in red and blue. The blue discs represent the 18 targets in the environment. When a target is foraged by the two agents, a red cross (resp. blue) is drawn on the target if the red agent (resp. blue) arrived on it first. Each snapshot corresponds to a trial where agents adopted a different strategy: (A) *turning* or (B) *leader/follower*.

In the turning strategy, both individuals turn around one another so that they can keep the other individual in their line of sight and stay close to it (see Figure 7.1(A)). At the same time, the two individuals try to get closer to a target. This way, as soon as one of the two individuals is in contact with a target, the other individual can join it so the target may be collected cooperatively. Consequently, both individuals adopt a similar behaviour in this strategy and can be described as generalists.

In the leader/follower strategy, the individuals specialise in two roles: a leader and a follower. The leader always gets on the target first and checks rarely for its partner. In comparison, the follower tries to keep its leader in view during the entirety of the simulation so that it can get on the same target (see Figure 7.1(B)). Consequently, we observe

the expression of two clearly heterogeneous behaviours which implies that both individuals are specialists. More importantly we also showed that, given our agents' capabilities, each phenotype needed to be encoded by a different genotype for specialisation to happen.

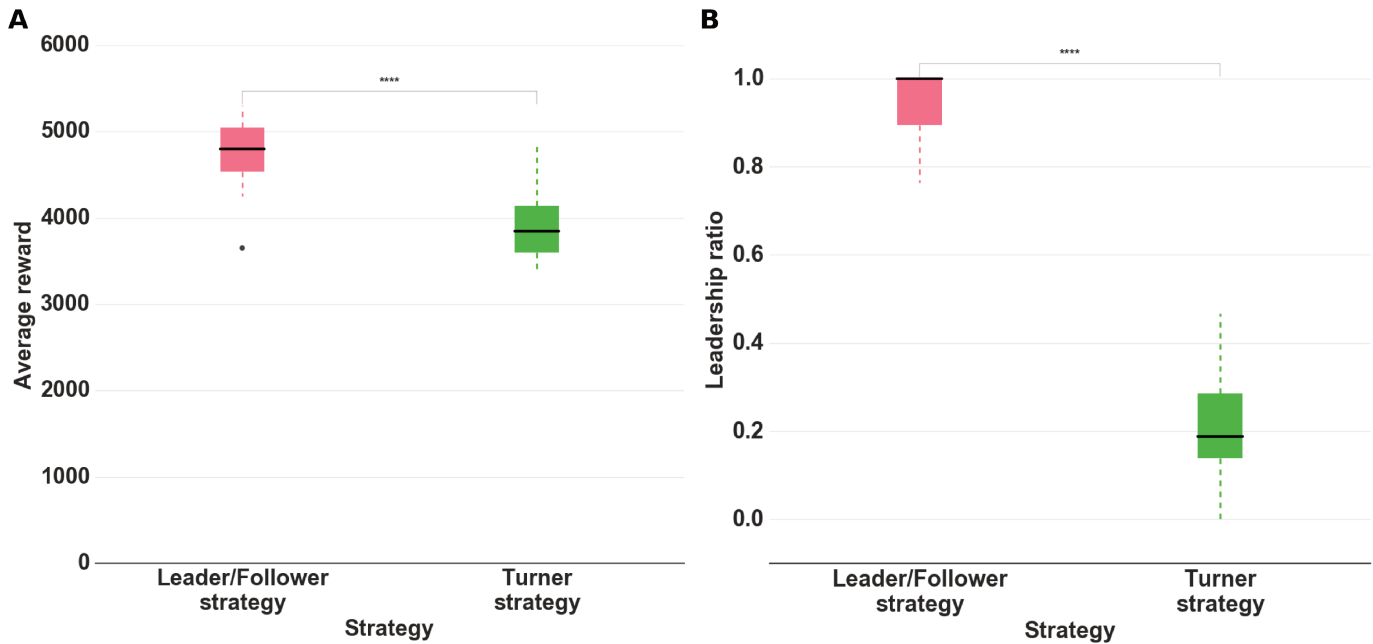


Figure 7.2 – Average reward and leadership proportion with a leader/follower or turning strategy Boxplots of (A) the average reward and (B) the leadership proportion over 20 independent trials for the leader/follower and turning strategies. The leadership ratio of an individual represents the propensity for one individual among the pair to arrive first more often than its partner on a target collected in a cooperative fashion. The position of each target at the beginning of each trial was randomized.

Figure 7.2(A) shows the efficiency of each strategy, defined as the average reward obtained by the two individuals during a simulation over 20 independent trials (with randomized targets' positions for each trial). We can see that, as expected, the leader/follower strategy achieves a significantly higher efficiency (Mann-Whitney U-test on the average reward over 20 trials, p -value < 0.0001). This difference in efficiency is directly correlated to a highly significant difference in the proportion of leadership as shown in Figure 7.2(B) (Mann-Whitney U-test on the leadership proportion over 20 trials, p -value < 0.0001). We compute this proportion by looking at the propensity for one of the two individuals to arrive first more often on a target foraged cooperatively (i.e. the emergence of a leader).

Evolving Heterogeneous Behaviours with an Elitist Selection

Bootstrapping leader/follower strategies

In this first experiment, we are interested in the emergence of a leader/follower strategy when starting with a population of random individuals under an $(\mu + \lambda)$ elitist selection. In order to investigate the influence of population size, we tested three different sizes N : 20, 40 and 100. For each population size, we conducted 11 independent runs, each one lasting 90000 evaluations. For each population size N , we defined μ (i.e. the number of parents) and λ (i.e. the number of offsprings) as $\frac{N}{2}$. For example, when population size was 100, 50 individuals were kept from the previous generation and used to create 50 mutated offsprings.

Pop. size	# L/F Strat.	# Turning Strat.	# NC Strat.	Total
20	0	11	0	11
40	0	11	0	11
100	1	10	0	11

Table 7.2 – *Strategies evolved by the best individuals under elitist selection with an initially random population.* Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the replicates for different population sizes N . We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under an $(\mu + \lambda)$ elitist selection, with $\mu = \frac{N}{2}$ and $\lambda = \frac{N}{2}$. Individuals' genotype values were initially random. In the table "L/F" stands for leader/follower and "NC" for "Non-Cooperative".

Table 7.2 shows the repartition of the best individuals' strategies at the last generation of evolution for each population size. We consider a behaviour to be cooperative when more than 50% of the total number of targets collected are foraged cooperatively. First, we observe that in every replicate individuals always end up evolving a cooperative strategy. We also see that evolving a leader/follower strategy is difficult as specialists evolve in only 1 run (out of 33) and when the population size is 100. These results suggest that it is nearly impossible to evolve such heterogeneous behaviours with this setting.

However, when looking at the whole evolutionary history we can reveal additional information about the evolution of specialists. We show in Figure 7.3 the proportion of evolutionary time when the best individual of each run adopted a leader/follower strategy. This value is computed as the ratio of the number of generations when the leadership ratio was high enough (over a threshold value of 0.6) out of the total number of generations. We observe that even if the best individuals end up adopting a generalist strategy, this was not the case during the entirety of the evolution. In particular, there is a significant increase (Mann-Whitney, p -value < 0.05) in the number of generations where the best individual showed a leader/follower strategy when population size was 100 compared to

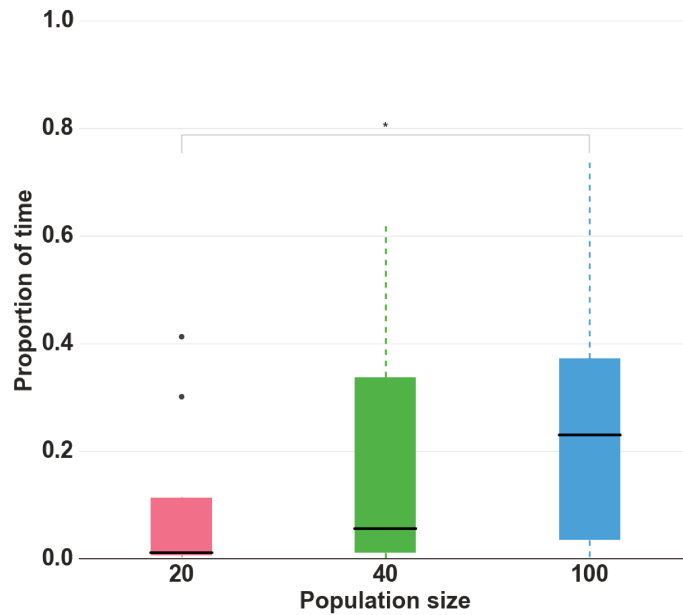


Figure 7.3 – Proportion of time with a leader/follower strategy. Boxplots of the number of generations where the best individual in each replicate adopted a leader/follower strategy out of the total number of generations. We consider that the best individual adopted a leader/follower strategy when its leadership ratio was over a threshold value of 0.6.

a population size of 20. Therefore this implies that it is possible to evolve specialists but their stability in the population over time is nearly impossible to achieve.

Maintaining heterogeneity in a population seeded with specialists

In order to investigate the lack of stability of genotypic polymorphism under elitist selection, we design another experiment. We separately evolve a population of efficient *leader individuals* and *follower individuals* beforehand. We then replace the worst individuals w.r.t. fitness score in the population of leaders by a certain amount of followers. Our goal is to study if artificially constructing such population could result in the invasion and fixation of a stable leader/follower strategy.

The number of followers initially inserted in the population was varied according to two different settings: (1) we add only one follower or (2) we add an amount of followers equal to half of the population. Experiments were replicated 11 times during 90000 evaluations with population size of 40 and 100.

We show (Table 7.2) no significant differences in comparison to simulations with a population constituted of initially random individuals w.r.t. the number of simulations where a leader/follower strategy evolved. These results suggest that even when purposely adding specialists, their stability in the population is still very hard to achieve. This implies that whether the behaviours are evolved from random genotypes or bootstrapped with efficient individuals is not as important as maintaining heterogeneity in the population. In particular, in only one replicate among the 3 runs where a leader/strategy was

Pop. size	Followers added	# L/F Strat.	# Turning Strat.	# NC Strat.	Total
40	1	0	11	0	11
40	20	0	11	0	11
100	1	1	10	0	11
100	50	2	9	0	11

Table 7.3 – Strategies evolved by the best individuals under elitist selection when adding followers. Repartition of the different strategies adopted by the best individuals at last evaluation in each of the replicates for different population sizes N . We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under a $(\mu + \lambda)$ elitist selection, with $\mu = \frac{N}{2}$ and $\lambda = \frac{N}{2}$. The population was initially seeded with a population of leaders in which we added a specific amount of followers. In the table "L/F" stands for leader/follower and "NC" for "Non-Cooperative".

eventually adopted (out of 44) did the specialists initially added were maintained. In the 2 other runs we observe multiple emergences and disappearances of specialists throughout evolution.

Evolution Under a Fitness-Proportionate Selection

In this next experiment we want to investigate the evolution of heterogeneous behaviours when using a fitness-proportionate selection. As fitness-proportionate is known to allow frequency-dependent selection, we hypothesize that it may facilitate the evolution of specialists.

Bootstrapping leader/follower strategies

Similarly to the elitist selection, we replicated our experiments in 11 independent runs during 90000 evaluations. Likewise, population sizes were 20, 40 and 100.

We show in Table 7.4 that results are highly different when using such selection scheme. In particular, the fitness-proportionate selection performed poorly w.r.t. evolving cooperative strategies. For each population size, no cooperative strategy evolved at all in the vast majority of replicates. However in one particular run we do observe the emergence and fixation of specialists. This is similar to what was observed under elitist selection w.r.t. evolving specialists.

Yet a closer look at the dynamics of evolution under a fitness-proportionate selection yields interesting results. In particular, there is not much variation in the strategy adopted by the best individuals throughout evolution. This is consistent with the fact that the bootstrap of a cooperative strategy was not observed in most of the replicates: fitness-proportionate is not efficient in evolving any cooperative behaviour. In consequence, there is not much variation in the proportion of individuals adopting a leader/follower strategy during evolution. As a matter of fact, we observe that in the only replicate where there was

Pop. size	# L/F Strat.	# Turning Strat.	# NC Strat.	Total
20	0	1	10	11
40	0	1	10	11
100	1	2	8	11

Table 7.4 – Strategies evolved by the best individuals under fitness-proportionate selection with an initially random population. Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the replicates for different population sizes. We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under a fitness-proportionate selection. Individuals’ genotype values were initially random. In the table “L/F” stands for leader/follower and “NC” for “Non-Cooperative”.

genotypic polymorphism at the end of the simulation, specialists were already present at the random initialisation of the population and did not evolve through mutation. This is very different with the elitist selection where we observe multiple emergences of specialists (even briefly) during evolution in many different runs.

Maintaining heterogeneity in a population seeded with specialists

Pop. size	Followers added	# L/F Strat.	# Turning Strat.	# NC Strat.	Total
40	1	7	0	4	11
40	20	8	0	3	11
100	1	10	0	1	11
100	50	10	0	1	11

Table 7.5 – Strategies evolved by the best individuals under fitness-proportionate selection when adding followers. Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the replicates for different population sizes N . We indicate in each cell the number of simulations where a particular strategy evolved. Populations were evolved under a fitness-proportionate selection. The population was initially seeded with a population of leaders in which we added a specific amount of followers. In the table “L/F” stands for leader/follower and “NC” for “Non-Cooperative”.

As expected from previous results, fitness-proportionate performs well in terms of stability of heterogeneous behaviours. We show in Table 7.5 that in the majority of replicates the best individuals adopt a leader/follower strategy at the end of the simulations. This is particularly true when population size is high enough (100). A major difference with the elitist selection is that in all replicates where a leader/follower strategy was observed at the end of the run, the specialists were maintained from the start throughout evolutionary time. These results suggest that, although not efficient at bootstrapping cooperative

behaviours, fitness-proportionate performs well w.r.t. the stability of genotypic heterogeneity. Furthermore, we can hypothesize that this selection scheme is good at maintaining heterogeneity specifically because it largely fails (under our choice of parameters) at bootstrapping any cooperative strategy.

Computational Analyses of Population Dynamics

In this present section, our goal is to understand more deeply the dynamics at play which allow the invasion of suboptimal generalists even when efficient specialists are present. To that end we run computational analyses based on the expected fitness of each of the three phenotypes. Table 7.6 shows the average payoff of pair-wise simulations between each type of phenotypes. We consider the payoffs for both phenotypes in each pair to be identical as no significant differences were observed between their payoffs.

Phenotype	Leader	Follower	Turner
Leader	1265	5000	3480
Follower	5000	100	2750
Turner	3480	2750	2755

Table 7.6 – Payoff matrix for pair-wise simulations of each phenotype. Average payoffs of each phenotype against every phenotype in a pair-wise simulation. Each pair was evaluated 10 times in order to decrease the stochastic effects of the initial conditions (i.e. random positions of the targets).

Several observations can be made directly from these results. First, we can confirm that the leader/follower strategy displayed by a (*leader*, *follower*) pair is clearly the best strategy. However each one of these two phenotypes performs very poorly against itself with the worst payoff obtained by a pair constituted of two *followers*. Secondly, *turner* individuals perform also very well against *leaders*. Last, there is no significant differences w.r.t. payoffs when a *turner* is paired with a *follower* or another *turner*. These last two points hint at a shared lineage between *followers* and *turners*.

Indeed analyses of the genotypes' histories in our previous experiments reveal that *turner* individuals in fact descend from *follower* individuals. This means that they act as *followers* when interacting with *leaders* but are not as efficient. However they are a lot more efficient than *followers* when paired with individuals of the same phenotype (or *followers*).

From this payoff matrix, we run computational analyses to model the gradient of phenotypes' repartition in an infinite population. The fitness W of a particular phenotype i is computed as follows:

$$W_i = \sum_{j=1}^M P(ij) * F(j)$$

with j the phenotype it is paired with, M the number of different phenotypes (3), $P(ij)$ the payoff of phenotype i against j and $F(j)$ the proportion of phenotype j in the population. From this fitness, we can deduce the variation of phenotypes repartition by updating the proportion F of each phenotype i :

$$F_i = F_i * \frac{W_i}{\sum_{j=1}^M W_j}$$

We show in Figure 7.4(A) a vector field of this gradient. We can see that there actually exists an equilibrium between the three phenotypes (marked by the a dot at the crossing between the the dotted lines). This implies that even though the *turner* strategy is not the more efficient one, it is still expected that this phenotype can invade and coexist with the two other phenotypes.

We can hypothesize that we could not observe this equilibrium in our robotic simulations because of the stochastic effects arising from selection in a finite population. In order to study this hypothesis we ran additional computational simulations based on the same payoff matrix. The initial population is entirely composed of *leaders* and the selection method is an elitist ($\frac{N}{2} + \frac{N}{2}$) evolution strategy where N is the population size. Every 10 generations, each offspring has a probability of $1 * 10^{-2}$ to mutate into any of the two other phenotypes.

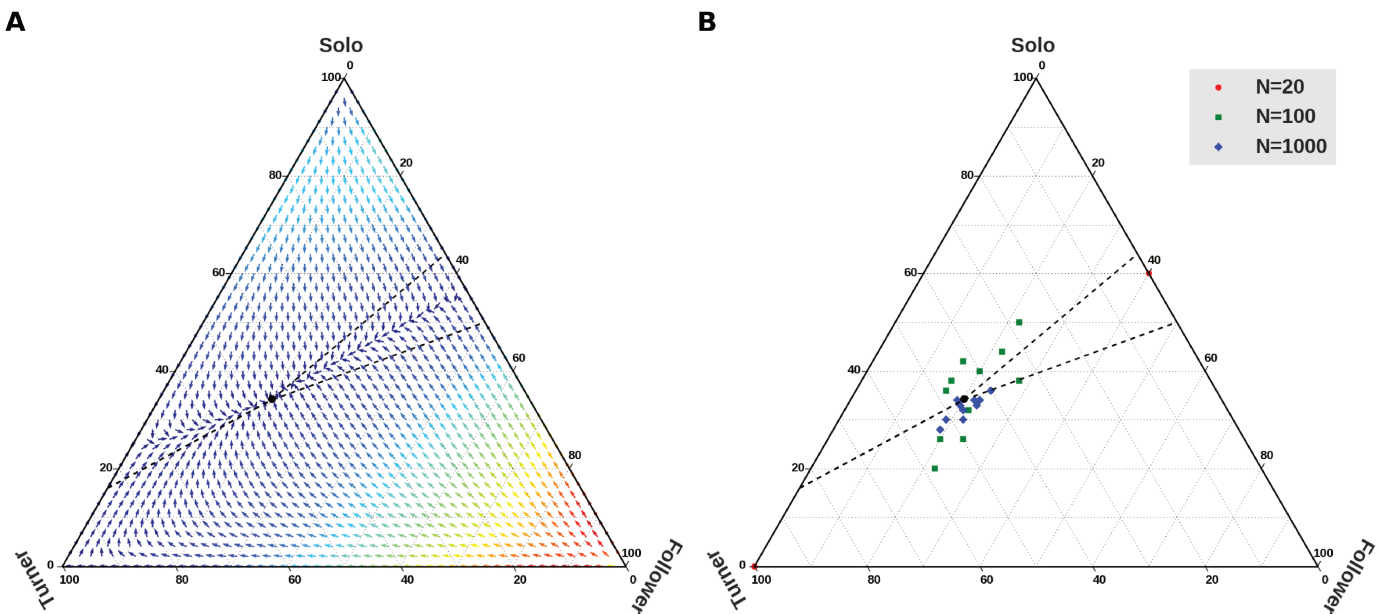


Figure 7.4 – Vector field of the gradient of phenotypes’ proportions and proportions of phenotypes at last generation of evolution. (A) Vector field of the gradient of phenotypes’ proportions in an infinite population. The strength of variation is indicated by the color of the arrow. (B) Repartition of phenotypes at the last generation of evolution for all three population sizes. Evolution lasted 1500 generations and results were replicated across 11 independent simulations. The initial population was entirely composed of leaders.

Figure 7.4(B) shows the final repartition of phenotypes after 1500 generations of evolution for $N = 20$, $N = 100$ and $N = 1000$ in 11 independent replicates. We can see that when increasing population size we also increase the probability that an equilibrium where the three phenotypes exist is reached. We actually observe that the repartition of phenotypes at last generation of evolution gets closer to the predicted equilibrium as population size increases. This implies that when population size increases, the probability to lose particular phenotypes decreases. In other words, the effect that the stochasticity of fitness evaluation has on the sampling of the genotypes for the next generation is mitigated: population size is essential to the maintenance of specialists.

Key Properties for Evolving Heterogeneous Behaviours

From the previous Section, we can hypothesize two key properties for the successful evolution of genotypic polymorphism. First, we showed that population size needed to be large enough in order to decrease the probability that heterogeneity could be lost during the evolutionary time. Even under an elitist selection where the best individuals are immediately selected, the stochastic nature of fitness evaluation entails that there is no guarantee that both types get selected. This means that a performance biased selection may lead to the composition of the new population not accurately representing the genotypic diversity of the previous one. Therefore, there needs to be a mechanism for the preservation of genotypic diversity. Second, we previously saw that one key reason for the invasion of *turner* individuals is that, while *followers* perform badly against themselves, this is not the case for the formers. This means that the manner in which robots are paired is essential for achieving specialisation.

In order to test these hypotheses we design a last experiment where we diverge from the initial problem and now coevolve two separate populations. In this coevolution algorithm, each individual of one population is always evaluated against an individual of the other population (5 times as in previous experiments). Then, each population separately undergoes selection under an elitist (10 + 10) selection method to create the population of the next generation (which means that each population size is 20). We conducted 11 independent replicates which lasted 90000 evaluations each. The populations were initially constituted of random individuals.

# L/F Strat.	# Turning Strat.	# NC Strat.	Total
11	0	0	11

Table 7.7 – Strategies evolved by the best individuals when coevolving two populations. *Repartition of the different strategies adopted by the best individuals at the last evaluation in each of the 11 replicates. We indicate in each cell the number of simulations where a particular strategy evolved. Two populations were coevolved under elitist selection and the individuals' genotype values were initially random. In the table "L/F" stands for leader/follower and "NC" for "Non-Cooperative".*

We show (Table 7.7) that when using coevolution, we always evolve specialists in every replicates. Moreover, this algorithm is highly stable as the heterogeneous behaviours that emerged were never lost during evolution in every replicates. This means that coevolution is highly efficient both for the bootstrap of a leader/follower strategy and its maintenance throughout evolution. Regarding our hypothesized properties, we can check that the coevolution algorithm respects both of them. Firstly, as populations are separately coevolved, we make sure that performance-based selection does not accidentally lead to the disappearance of specialists. Thus we ensure that the populations' genotypic diversity is highly protected. Secondly, we create a very specific pairing between individuals. Indeed individuals inside the same population are never partnered with one another. This means that *followers* are always paired with *leaders*. As *turners* thus possesses no fitness benefit over the other phenotypes, their invasion is prevented. The question is open as to how to endow an algorithm working on a single population with such properties.

Discussion and Conclusions

In this Chapter, we investigated the evolution of specialisation through a leader/follower strategy in a cooperative foraging task. Our goal was to reveal the difficulties that arise when trying to evolve genotypic polymorphism in a single population. To that end, we mainly studied the dynamics of evolution with two different selection methods: an $(\mu + \lambda)$ elitist evolution strategy and fitness-proportionate selection.

We first showed that the long term evolution of a leader/follower strategy was nearly impossible with an elitist selection. However bootstrapping specialists was not a problem as we observed that they frequently emerged during evolution. The major obstacle was rather to maintain heterogeneity over evolutionary time. Indeed, even when adding efficient followers to a population of leaders to force the adoption of a leader/follower strategy, specialists couldn't be maintained. In comparison, the properties shown by the fitness-proportionate algorithm were quite the opposite. While it was almost not capable of evolving a leader/follower strategy (nor any other cooperative strategy), the fitness-proportionate selection demonstrated high stability. It was therefore capable of maintaining specialists when present. We thus revealed two critical properties for evolving heterogeneous behaviours in a single population: *bootstrapping* these behaviours and *maintaining* them throughout evolution.

We then ran computational analyses and showed that while a pair of turners is indeed less efficient w.r.t. payoff than a pair of leader and follower, it is a lot more efficient than a pair of leaders or a pair of followers. As a result, these individuals can easily invade part of the population. Moreover, we also showed that the maintenance of specialists was very sensible to population size. Performance biased selection can indeed affect heterogeneity in the composition of the next generation's population. Finally, a coevolution algorithm, which we showed to be always successful in evolving heterogeneous behaviours, solved both of these two problems with (1) *specific partners selection* as pairs were constituted of individuals from different populations and (2) *protection* of the behaviours evolved by applying selection separately on the two populations. While this algorithm is not

concerned with genotypic polymorphism in a single population, it is useful to yield effective mechanisms which could be studied to solve our problem.

This raises several interesting perspectives on how to solve this problem. First, niche protection could prevent the disappearance of the efficient but unstable leader/follower strategy. As a matter of fact, coevolution is akin to a particular type of niches protection with 2 niches. However, such mechanism could be implemented without specifying the explicit number nor the organization of the niches. Rewarding diversity (Lehman and Stanley 2008) is also known as an effective way to protect novel behaviours and could be another promising direction. In particular, a multiobjective algorithm on performance and diversity (Doncieux and Mouret 2014), by rewarding genotypic and phenotypic diversity, may protect evolved specialists.

Secondly, we showed that because partners were chosen randomly among the population, it created the opportunity for a "parasitic" strategy to invade. An interesting direction for future works could be to investigate restrictions in the choice of partners. For example it would be compelling to investigate how the individuals could evolve to select their partner based on genotypic or phenotypic information.

Chapter 8

Discussion

8.1 Modeling the Evolution of Cooperation	139
8.1.1 Summary of Contributions	139
8.1.2 Limits and Perspectives	141
8.2 Automatic Design of Collective Robots	143
8.2.1 Summary of Contributions	143
8.2.2 Limits and Perspectives	145
8.3 Concluding Remarks	146

The contributions of this thesis have been divided in two parts that, while they revolve around evolutionary robotics, are centered on very different problems. As such, we choose to discuss and conclude on each part separately.

8.1 Modeling the Evolution of Cooperation

8.1.1 Summary of Contributions

In the first Part of this manuscript we have been interested in the influence of coordination on the evolution of mutualistic cooperation. We focused in particular on the proximate explanations of coordination and their impact on the evolution of cooperation. Most studies centered on cooperation have often been dedicated to explaining the stability of altruistic behaviours. Mutualistic actions in comparison have usually been ignored. However, while mutually beneficial actions do not raise any issue of stability, the origin of this type of cooperative behaviours is not trivial. Because they require coordination, the spread of a cooperative behaviour from an initial population of solitary individuals remains an open question. We thus were interested in the bootstrapping problem faced by the emergence of mutually beneficial cooperation.

To that end, we used evolutionary robotics as our modeling framework. Because we were interested in the origin of cooperative actions rather than their stability, the modeling of mechanistic constraints is critical. Namely, the convergence to a cooperative solution



is impacted by the availability of mutations, i.e. the possibility for a particular mutant to appear in the population. While classical models in evolutionary biology are of great help in understanding the dynamics of the evolution of cooperation, they make critical assumptions with regard to the convergence towards a cooperative solution. This is what has motivated our choice to use individual-based modeling and evolutionary robotics in particular as a modeling method. In particular, evolutionary robotics allows to model the mapping between genotype and phenotype and thus study the proximate mechanisms at play in the evolution of cooperation.

In a first study, we focused on the impact of modeling the mechanics of behaviour in the evolution of mutualistic cooperation. To that end we took inspiration from the game theoretical model of the stag hunt and studied the transition from the solitary equilibrium (hare hunting) to the cooperative equilibrium (stag hunting). We first revealed that there is a drastic difference between the results predicted by classical models in game theory and what we observed in evolutionary robotics. With a classical model, the transition to stag hunting always occurred. In comparison, with a model in evolutionary robotics, this transition was nearly impossible. Furthermore, we showed that even under maximal genetic relatedness (i.e. individuals were clones of each other), the evolution of cooperation was still unlikely. We thus revealed that the mechanistic constraints are critical for the origin of cooperation in the stag hunt game. The evolution of cooperation is faced with a chicken and egg dilemma: for cooperation to be selected, it needs to be beneficial. Yet the success of the cooperative action requires that others have evolved the capacity to coordinate, which is not beneficial on its own. If one assumes that a single mutation can lead to the evolution of cooperation then it is considered that the same mutation is responsible for both the modification of the preferred prey and the capacity to coordinate. We showed here that doing so hides part of the issue. Moreover we showed in our simulations that the individuals needed to evolve a complex behaviour in order to be able to coordinate. Therefore, it is necessary to take into account the mechanics of behaviours in order to fully understand the evolution of cooperation. This means that there is a need for complementary frameworks which model these mechanisms.

In the second Chapter, we studied how the nature of coordination behaviours may impact the transition between collective equilibria. More precisely, we focused on the issue of selecting between multiple stable equilibria. When a collective equilibrium has emerged, no single mutant has a selective advantage to deviate from this equilibrium. As such, this raises the issue of the transition to the optimal equilibrium. Our goal was to study if individual selection alone could lead to the optimization of group-traits. To that end, we used a model of collective hunting where individuals could choose between two types of prey: boar (suboptimal prey) and stag (optimal prey). We revealed that the transition towards the optimal prey was impossible under simple environmental features (only one prey of each type). However, we revealed that results were different when more realistic assumptions about collective hunting were made. Surprisingly, when the environment was more complex (nine prey of each type), the switch to stag hunting could sometimes occur. Under such environmental conditions, it was necessary to collectively decide on which prey to hunt. This meant that the evolution of coordination was required in order to achieve cooperation. In consequence, each individual evolved the capacity

to react to the other individual's behaviour. Thus a mutant could indirectly change the behaviour of the group and thus lead to stag hunting. However, we observed that the coordination strategy evolved was a very symmetrical one where both individuals could adopt the same behaviour. We then increased the neural complexity of individuals and showed that the transition to stag hunting was highly facilitated. Furthermore, we revealed that individuals evolved a strongly asymmetrical coordination strategy through specialisation: a leader/follower strategy. In this strategy, the follower only reacted to the behaviour of the leader which it tried to follow. This meant that a mutation on the leader was sufficient to change the group's behaviour and thus reap the benefits of cooperative stag hunting. Additionally, we observed that this strategy was more efficient than the previous one in terms of rewards obtained. We thus showed that the evolution of an individually adaptive strategy led to the transition to the optimal collective behaviour. Therefore, it was possible for individual selection alone to explain the optimization of group traits.

In both of these studies, we thus revealed the critical role of coordination in the evolution of cooperation. In consequence, we demonstrated that it is indeed necessary to take the mechanics of coordination behaviours into account in order not to neglect crucial aspects of the evolutionary dynamics. Additionally, we also presented a general mechanism for the evolution of cooperation in the stag hunt. The boar we introduced in our study on the selection of equilibria could act as an evolutionary pathway in the stag hunt. Namely, this prey can be hunted alone but rewards more when it is hunted cooperatively, which is a realistic expectation for hunting in the natural world. As such, coordination and cooperation could initially be bootstrapped on this prey (because it is not as risky as hunting stags). Then, we showed that when coordination was present it was possible for individual selection to optimize the collective behaviour. In consequence, this could lead to the transition to the purely cooperative equilibrium: stag hunting.

8.1.2 Limits and Perspectives

The evolution of communication

During this thesis, we have studied coordination strategies that did not require any direct communication between individuals. We wanted to study if coordination was possible without endowing individuals with communication capabilities. Indeed, we aimed at keeping the complexity of our robot model as low as possible so that we could focus on the most basic mechanisms that could lead to the evolution of cooperation. More complex agents could beg the question of the role of the robots' capabilities on the observed phenomena. In particular, the goal behind the first Part of our study was to compare classical models in evolutionary biology with modeling under an evolutionary robotics framework. As such, it was important that no particular design choice could alter the relevance of this comparison.

We found that our robots were capable of coordination without any means of communication. Indeed, they could evolve a surprising behaviour which, while it was not the most efficient way to coordinate, allowed them to cooperate. In a way, we can make here

a similar observation as Mitri and colleagues (Mitri et al. 2009). Namely, individuals could rely on indirect communication cues resulting from their embodiment. However we could also implement a more direct way for them to communicate. Communication is used in numerous different social species, among which we already gave the example of the spotted hyenas (Drea and Carter 2009; Smith et al. 2010, 2012). They are capable of using signaling techniques and communication in order to achieve high level of coordination during collective hunting. As such, while the emphasis of our second study was mainly put on division of labour, communication may deeply impact the nature of coordination behaviours. In consequence, we hypothesize that the evolution of communication strategies could affect the evolution of collective actions. Additionally, an interesting perspective would be to let the individuals evolve how to communicate from basic communication capabilities (e.g. the broadcast of a simple signal). As such, this could lead to coevolutionary dynamics between communication strategies and coordination behaviours.

Moreover, the evolution of communication has already been studied in several works in evolutionary robotics. We already mentioned in Chapter 2 that some have been interested in the general evolution of communication between foraging robots (Floreano et al. 2007), information suppression (Mitri et al. 2009) or the role of historical contingencies on the evolution of signaling strategies (Wischmann et al. 2012). As with the evolution of cooperation, few have been interested in the emergence of communication among unrelated individuals. A notable counterexample is the work of Solomon et al. (Solomon et al. 2012) who modeled communication strategies between hyenas. However, the impact of communication on the origin of mutualistic actions has not been studied.

Cooperation between bigger groups of agents

Our initial inspiration was the game theoretical framework of the stag hunt. As such interactions take place between only a pair of individuals. However, it could be interesting to increase the number of agents and to study how the evolutionary dynamics would change. In particular, in our second study we have been interested in the optimization of group-traits by way of individual selection. We showed that the transition between multiple equilibria could occur without any group-level mechanism. As we provided no evidence that our demonstration could be scaled up to more than two individuals, it could be argued that what we consider a group-trait is limited in this context.

However, we believe that similar results would be observed in larger groups of individuals. We even hypothesize that the results could be more explicit in this case. Assume that we have a scaled-up version of our study where interactions take place between 5 agents. Hunting stag would now require that the 5 individuals cooperate together to reap the benefits of the hunt. Because the coordination of 5 individuals is more challenging than that of a pair, the evolution of efficient coordination strategies would be even more favored. As such, the evolution of asymmetrical behaviours like the leader/follower strategy should have a bigger impact on the transition towards the optimum. Indeed, it would be impossible for a single mutant to lead the group to stag hunting without an asymmetrical strategy. We have begun working on preliminary experiments where we increase the number of agents in our simulation (with groups of 3, 4 or 5 robots).

Online evolution

An interesting continuation of our work would be to model the evolution of mutualistic cooperation in a more ecologically realistic setup. More precisely, in the classical framework of evolutionary robotics we use an offline evolutionary algorithm. This initially means that there is a separation between the design of a robot and its deployment; robots evolve in a different environment than the one where they are used. In comparison, under an online paradigm evolution occurs directly in the operational environment. The field of embodied evolution (Watson et al. 2002) was created in order to address this question and in particular the issue of transferability that stems from evolving robots in an offline manner. As such, embodied evolution is mainly concerned with design questions. In our case, it would be interesting to use online evolution as another manner in which to model evolutionary phenomena.

Furthermore, in the particular case of *environment-driven* embodied evolution the selection pressure is driven by the environment. In classical ER (and embodied evolution), a fitness function is used to evaluate the performance of individuals and determine if they will produce offspring. In this case, there is an objective-driven selective pressure. This is different from the biological definition of fitness, where fitness is an *a posteriori* evaluation of the reproductive success of a given individual. As such, a more realistic approach to the modeling of evolution would be to require that the individuals meet with each other in order to exchange genetic material (Bredeche and Montanier 2010). This is the principle of environment-driven embodied evolution. In consequence, a perspective would be to study the evolution of mutualistic cooperation under such paradigm. In this case an interesting feature is that the selection process also evolves.

Few have focused on this aspect of biological modeling. Montanier and Bredeche investigated the evolution of altruism among a population of simulated robots under an online environment-driven algorithm (Montanier and Bredeche 2011, 2013). This way they could study the impact of genetic relatedness and dispersion on the emergence and stability of altruistic behaviours.

8.2 Automatic Design of Collective Robots

8.2.1 Summary of Contributions

In the second Part of this thesis our goal was to study how to design the evolution of cooperation in evolutionary robotics. More precisely, we were interested in the impact of genetic team composition on the evolution of efficient coordination strategies in multirobot systems. Multirobot systems have multiple advantages in comparison to using a single robot among which robustness, efficiency and the capacity to achieve tasks that a single robot could not. However, because they require the control of several agents, they are also more challenging to design. In this context, multiple different architectures exist for multirobot systems and several ways to design the control of collective robots have been proposed. But while a popular and often efficient method has been to manually design the robots, there has been a strong interest in creating methods to automatically design

them. Automatic design can lead to robots which better react to environmental changes and unknown environment. Additionally, some problems are simply nearly impossible to design in a manual fashion. In this context, we focused on evolutionary robotics as an approach to the automatic design of distributed robots.

An open issue when designing cooperative robots in evolutionary robotics is team composition. Namely, the robots that constitute a team can be homogeneous or heterogeneous, whether in terms of morphology or control. Here we focused on the influence of team composition on the control of agents only. The classical approach in evolutionary robotics has been to use an homogeneous group of robots, where every individual comes from the same genotype. However, it is argued that heterogeneity could lead to more diverse behaviours between robots and thus generate higher efficiency.

In the first Chapter of this Part, we compared homogeneous and heterogeneous approaches on two criteria: evolvability and efficiency. We designed a collective foraging task inspired by the stag hunt where the individuals could forage two types of resources: one that could be foraged alone and an other more rewarding that needed to be collected cooperatively. In this context, we use a restricted definition of evolvability where it corresponds to the capacity of a particular method to evolve cooperators. In comparison, efficiency refers to the performance of the cooperative solution w.r.t. resources collected. We compared a clonal approach (i.e. where individuals are homogeneous) to two aclonal ones (i.e. where the composition is heterogeneous): one where individuals were taken from the same population and the other where they came from two separately coevolved populations. We revealed that there was a tradeoff between evolvability, which is best achieved with the clonal approach and efficiency, where coevolution evolved more efficient cooperative strategies. In particular we showed that division of labour would systematically evolve with coevolution. In order to go beyond this tradeoff and improve each method on both criteria we then added incremental evolution. The goal of incremental evolution is to decompose a complex task into several sub-tasks that are evolved separately in order to ease the learning process. In our case, we pre-evolved our individuals in a simpler cooperative task. We showed that while this produced no significant differences for the clonal approach, the evolvability of coevolution was greatly increased. In consequence we showed that an aclonal approach, coevolution, was the best method on both evolvability and efficiency. However, this increase in evolvability comes at the price of a pre-evolution step. We thus revealed a new tradeoff: coevolution may outperform a clonal approach but at the cost of additional computations.

In the next Chapter, we focused on the evolution of specialisation between heterogeneous individuals. We took a simpler task than that of the previous Chapter where cooperation is easy to evolve but efficient coordination strategies are favoured. In this task, we showed that two coordination strategies could evolve: a turner strategy where both individuals are generalists and a leader/follower strategy where the two robots specialise. We thus wanted to study how division of labour could evolve between heterogeneous individuals at the level of the population. Namely, we studied the evolution of genotypic polymorphism, i.e. the coexistence of several different genotypes (encoding for diverse phenotypes) in a single population. We compared two selection schemes based on their capacity to achieve genotypic polymorphism: a $(\mu + \lambda)$ evolution strategy and

fitness proportionate. We revealed that while specialists could easily evolve under an elitist selection they could rarely be maintained throughout evolution. In comparison, fitness proportionate easily maintained genotypic polymorphism but was not efficient at evolving specialists. In order to understand the evolutionary dynamics at play, we then ran computational analyses based on the expected fitness of each strategy (turner, leader or follower) against every other strategy. We showed that generalists could invade the population by benefiting from the fact that specialists needed to be paired with complementary specialists in order to be efficient. Additionally, we revealed that under small population sizes, genetic diversity in the population could be lost during selection thus leading to the disappearance of specialists. From these results, we extracted two key properties for the evolution of genotypic polymorphism: stability of genotypic diversity and protection against the invasion of cheaters. While we could not achieve genotypic polymorphism in our study, we argued that an algorithm validating these properties could achieve this goal.

8.2.2 Limits and Perspectives

Diversity and novelty

A popular open issue in evolutionary robotics is about the selective pressures of evolutionary algorithms (Doncieux et al. 2015). The view of evolution as an optimization process led to the majority of works in ER typically relying on using a performance-based fitness. This means that an evaluation of performance is used to drive the search process toward the desired solutions. However, it has been shown that this approach of ER may lead to premature convergence which restricts the range of behaviours evolved as well as select solutions which are efficient on the short term only (Mouret and Doncieux 2012). In comparison there has been a recent interest for methods that do not rely exclusively on performance. Lehman and Stanley (Lehman and Stanley 2011) introduced *novelty search* for searching the goal of a maze and the evolution of bipedal walk. Selection was based on the novelty of behaviours compared to previous solutions rather than performance. They showed that this led to better results thanks to a more extensive research through the space of behaviours than with performance-oriented fitness. Mouret & Doncieux (Mouret and Doncieux 2012) used a multi-objective approach to optimize on both performance and behavioural diversity. They revealed that this allowed to improve performance in comparison to a more classical performance-based search.

In our case, we showed that there is a tradeoff between evolvability and efficiency. Aclonal approaches could evolve more efficient cooperative solutions but less easily than a clonal approach. Performance-based selection may in this case drive the evolutionary process to prematurely lose cooperative individuals as they are not efficient on their own. As such, multi-objective optimization on both diversity and performance could allow to maintain these individuals in the population and thus lead to the evolution of cooperation. In our study on genotypic polymorphism, diversity could also be used in order to protect the evolution of specialists. In this case we expect a population of specialists to be more diverse (in terms of genotype and phenotype) than one constituted of generalists. As such

diversity could allow to maintain the presence of division of labour in the population.

Preliminary results of multi-objective optimization with diversity did not unfortunately produce satisfying results w.r.t. evolving cooperation. Diversity is not a magic trick that luckily happens to find novel solutions. One of the main challenges when using diversity is to find an efficient measure of behavioural distance. While we think that diversity could be helpful in our case, finding both a correct and sufficiently general measure of this distance in the case of cooperation is challenging.

The transfer to real robots

In every experiment in this manuscript, we used simulated robots. While we believe our findings to mostly contribute to theoretical knowledge and not to be constrained by a particular robot model, we are aware of the open issue of transferability to real robots (i.e. the reality gap) (Jakobi et al. 1995; Mouret et al. 2012; Doncieux et al. 2015). As we already mentioned, several assumptions are made when using simulations of robotic agents. While we model a sensory system for our robots, sensors in the real world are noisy. Also, frictions can alter the way individuals move or collide with each other. As such, this creates a reality gap where evolved behaviours may not perform well when embedded in physical robots.

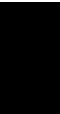
During this thesis, we also worked on building a robotic platform for the control of collective robots. The goal was to use simple and cheap robots in order to address distributed robotics questions. We thus designed robots composed of a Thymio-II, a raspberry PI and a camera module for raspberry. The Thymio constitutes the base of the robot and is equipped with two wheels and a collection of proximity sensors. The raspberry acts as the controller and can be used to write instructions to the Thymio and read sensory inputs.

As such one perspective could potentially be to apply our evolved behaviours in these real robots. This point might relate to both Parts of the thesis. However, in the case of modeling the evolution of cooperation, we would not learn more by using real robots rather than simulated ones. As previously explained, this would mostly ensure that no unrealistic physical assumptions are made in our simulations. We thus believe real robots to be more of an interesting perspective in our study of automatic design in evolutionary robotics. Preliminary experiments on those robots showed that the evolved leader/follower strategy could transfer well in simple situations. Additional experiments are needed to really validate the transferability of our solutions.

8.3 Concluding Remarks

The present work aimed at contributing to the field of evolutionary robotics in two different ways: model and design (Trianni 2014; Doncieux et al. 2015). On the one hand, we modeled the evolution of mutualistic cooperation. Thanks to evolutionary robotics we could show that the proximate aspects of behaviours are critical in understanding the origin of mutually beneficial actions. On the other hand, we studied the automatic design

of controllers for distributed robotics. We revealed that there are advantages in using heterogeneous teams of individuals and that it may be useful not to always resort to the classical approach of evolving clonal controllers. As a conclusion, we believe that it is of utter importance that these two aspects of evolutionary robotics are clearly separated, as we did throughout this thesis. However we also believe that by studying a common topic but with different motivations, each aspect could nurture our general intuitions about the other aspect.



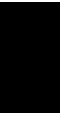
Publications

Articles

- Bernard, Arthur, Jean-Baptiste André, and Nicolas Bredeche (2015). “Evolution of Cooperation in Evolutionary Robotics : the Tradeoff between Evolvability and Efficiency”. In: *Proceedings of the European Conference on Artificial Life 2015*, pp. 495–502.
- (2016a). “Evolving Specialisation in a Population of Heterogeneous Robots: the Challenge of Bootstrapping and Maintaining Genotypic Polymorphism”. In: *Proceedings of the Artificial Life Conference 2016 (ALIFE XV)*, pp. 152–159.
- (2016b). “To Cooperate or Not to Cooperate: Why Behavioural Mechanisms Matter”. In: *PLOS Computational Biology* 12.5, pp. 1–14.

Invited talks

- Bernard, Arthur (2015). “Evolution of Cooperation in Evolutionary Robotics: the Tradeoff between Evolvability and Efficiency”. The 13th European Conference on Artificial Life (ECAL 2015). York, United Kingdom.
- (2016a). “Evolving Specialisation in a Population of Heterogeneous Robots: the Challenge of Bootstrapping and Maintaining Genotypic Polymorphism”. The 15th International Conference on the Synthesis and Simulation of Living Systems (ALIFE XV). Cancun, Mexico.
- (2016b). “Mechanistic Constraints in the Evolution of Coordination”. Journ’ee Evolution Artificielle Th’ematique (JET). Paris, France.



Bibliography

- Adami, Christoph (2012). “Boldly Going Beyond Mathematics”. In: *Science* 338.6113, pp. 1421–1422.
- Adami, Christoph, Jory Schossau, and Arend Hintze (2014). “Evolutionary Game Theory using Agent-based Methods”. In: pp. 1–19.
- Aiello, Leslie C and Peter Wheeler (1995). *The Expensive-Tissue Hypothesis: The Brain and the Digestive System in Human and Primate Evolution*.
- Albus, James S. (1991). “Outline for a theory of intelligence”. In: *IEEE Transactions on Systems, Man and Cybernetics* 21.3, pp. 473–509.
- Altenberg, Lee (1991). “Chaos from Linear Frequency-Dependent Selection”. In: *The American Naturalist* 138.1, p. 51.
- Alvard, Michael S (2003). “Kinship, Lineage, and an Evolutionary Perspective on Cooperative Hunting Groups in Indonesia”. In: *Human Nature* 14.2, pp. 129–163.
- Alvard, Michael S and David A Nolin (2002). “Rousseau’s Whale Hunt ?” In: *Current Anthropology* 43.4, pp. 533–559.
- Amato, Christopher, Girish Chowdhary, Alborz Geramifard, N Kemal Ure, and Mykel J Kochenderfer (2013). “Decentralized control of partially observable Markov decision processes”. In: *The 52nd IEEE Conference on Decision and Control (CDC)*.
- Arkin, Ronald C (1998). *Behavior-Based Robotics*.
- Åström, K.J (1965). “Optimal control of Markov processes with incomplete state information”. In: *Journal of Mathematical Analysis and Applications* 10.1, pp. 174–205.
- Axelrod, Robert (1984). *The Evolution of Cooperation*. New York: Basic Books, Inc.
- Axelrod, Robert and William D Hamilton (1981). “The Evolution of Cooperation”. In: *Science* 211.4489, pp. 1390–1396.
- Baldassarre, Gianluca, Stefano Nolfi, and Domenico Parisi (2003a). “Evolution of collective behavior in a team of physically linked robots”. In: *Applications of Evolutionary Computing* 2611, pp. 581–592.
- (2003b). “Evolving mobile robots able to display collective behaviors”. In: *Artificial life*.
- Baldassarre, Gianluca, Vito Trianni, Michael Bonani, Francesco Mondada, Marco Dorigo, and Stefano Nolfi (2007). “Self-organized coordinated motion in groups of physically connected robots”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 37.1, pp. 224–239.
- Banzhaf, Wolfgang, Guillaume Beslon, Steffen Christensen, James A Foster, François Képès, Virginie Lefort, Julian F Miller, Miroslav Radman, and Jeremy J Ramsden (2006). “From artificial evolution to computational evolution: a research agenda”. In: 7.September, pp. 729–735.

- Batut, B er enice, David P Parsons, Stephan Fischer, Guillaume Beslon, and Carole Knibbe (2013). “In silico experimental evolution: a tool to test evolutionary scenarios.” In: *BMC bioinformatics* 14.15, pp. 1–11.
- Bellman, Richard (1957). *A Markovian decision process*.
- Beni, Gerardo (2005). “From Swarm Intelligence to Swarm Robotics”. In: *Swarm Robotics*. Ed. by  ahin Erol and William M Spears. Vol. 3342. Berlin, Germany, pp. 1–9.
- Bergm uller, Ralph, Rufus a Johnstone, Andrew F Russell, and Redouan Bshary (2007). “Integrating cooperative breeding into theoretical concepts of cooperation.” In: *Behavioural processes* 76.2, pp. 61–72.
- Bernard, Arthur, Jean-Baptiste Andr e, and Nicolas Bredeche (2015). “Evolution of Cooperation in Evolutionary Robotics : the Tradeoff between Evolvability and Efficiency”. In: *Proceedings of the European Conference on Artificial Life 2015*, pp. 495–502.
- (2016a). “Evolving Specialisation in a Population of Heterogeneous Robots: the Challenge of Bootstrapping and Maintaining Genotypic Polymorphism”. In: *Proceedings of the Artificial Life Conference 2016 (ALIFE XV)*, pp. 152–159.
- (2016b). “To Cooperate or Not to Cooperate: Why Behavioural Mechanisms Matter”. In: *PLOS Computational Biology* 12.5, pp. 1–14.
- Bernstein, Daniel S, Robert Givan, Neil Immerman, and Shlomo Zilberstein (2002). “The complexity of decentralized control of Markov decision processes”. In: *Mathematics of Operations Research* 27.4, pp. 819–840.
- Beynier, Aur elie and Abdel-Allah Mouaddib (2011). “Applications of DEC-(PO)MDPs in Multi-Robot Systems”. In: *Decision Theory Models for Applications in Artificial Intelligence*, pp. 361–384.
- Binmore, Ken (2011). *Natural Justice*. Oxford University Press.
- Blumenthal, Joseph H and Gary B Parker (2004). “Co-Evolving Team Capture Strategies for Dissimilar Robots”. In: *AAAI Artificial Multi-Agent Learning Symposium*, pp. 15–23.
- Bonabeau, Eric, Marco Dorigo, and Guy Theraulaz (1999). *Swarm intelligence: from natural to artificial systems*, pp. 187–219.
- Bongard, Josh (2008). “Behavior Chaining : Incremental Behavior Integration for Evolutionary Robotics”. In: *in Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*. 1976. MIT Press, pp. 64–71.
- Bongard, Josh C. (2013). “Evolutionary Robotics”. In: *Communications ACM* 56.08, pp. 74–83.
- Botkin, Daniel B., James F. Janak, and James R. Wallis (1972). “Some Ecological Consequences of a Computer Model of Forest Growth”. In: *The Journal of Ecology* 60.3, pp. 849–872.
- Bourke, Andrew F G (2014). “Hamilton’s rule and the causes of social evolution.” In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 369.1642, p. 20130362.
- Bowling, Michael and Manuela Veloso (2003). “Simultaneous adversarial multi-robot learning”. In: *IJCAI International Joint Conference on Artificial Intelligence*, pp. 699–704.
- Boyd, Robert and Peter J. Richerson (1990). “Group selection among alternative evolutionarily stable strategies”. In: *Journal of Theoretical Biology* 145.3, pp. 331–342.
- Boyd, Robert and Peter J Richerson (2002). “Group beneficial norms can spread rapidly in a structured population.” In: *Journal of theoretical biology* 215.3, pp. 287–296.
- Bradtke, Steven J., Andrew G. Barto, and Pack Kaelbling (1996). “Linear least-squares algorithms for temporal difference learning”. In: *Machine Learning* 57, pp. 33–57.

- Brambilla, Manuele, Eliseo Ferrante, Mauro Birattari, and Marco Dorigo (2012). *Swarm robotics : A review from the swarm engineering perspective*. Tech. rep.
- Bredeche, Nicolas and Jean-Marc Montanier (2010). “Environment-driven Embodied Evolution in a Population of Autonomous Agents”. In:
- Brooks, Rodney A (1986). “A Robust Layered Control System for a Mobile Robot”. In: *IEEE Journal of Robotics and Automation* 2.1, pp. 14–23.
- Bshary, Redouan (2004). “Game Structures in Mutualistic Interactions: What Can the Evidence Tell Us About the Kind of Models We Need?” In: 34, pp. 59–101.
- Bshary, Redouan and Alexandra S Grutter (2005). “Punishment and partner switching cause cooperative behaviour in a cleaning mutualism.” In: *Biology letters* 1.4, pp. 396–399.
- Bshary, Redouan and Rui F Oliveira (2015). “Cooperation in animals: toward a game theory within the framework of social competence”. In: *Current Opinion in Behavioral Sciences* 3, pp. 31–37.
- Bullinger, Anke F, Emily Wyman, Alicia P Melis, and Michael Tomasello (2011). “Coordination of Chimpanzees (Pan troglodytes) in a Stag Hunt Game”. In: *International Journal of Primatology* 32.6, pp. 1296–1310.
- Calcott, Brett (2007). “The other cooperation problem: generating benefit”. In: *Biology & Philosophy* 23.2, pp. 179–203.
- Candea, Ciprian, Huosheng Hu, Luca Iocchi, Daniele Nardi, and Maurizio Piaggio (2001). “Coordination in multi-agent RoboCup teams”. In: *Robotics and Autonomous Systems* 36.2-3, pp. 67–86.
- Cao, Y.U., A.S. Fukunaga, A.B. Kahng, and F. Meng (1997). “Cooperative mobile robotics: antecedents and directions”. In: *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots* 23, pp. 226–234.
- Caraco, Thomas and Jerram L Brown (1986). “A game between communal breeders: When is food-sharing stable?” In: *Journal of Theoretical Biology* 118.4, pp. 379–393.
- Clutton-Brock, Tim (2002). “Breeding together: kin selection and mutualism in cooperative vertebrates.” In: *Science (New York, N.Y.)* 296.5565, pp. 69–72.
- Clutton-Brock, Tim, J M O’Riain, Peter N M Brotherton, D Gaynor, R Kansky, Ashleigh S Griffin, and M Manser (1999). “Selfish Sentinels in Cooperative Mammals”. In: *Science* 1 284, pp. 1640–1644.
- Connor, Richard C. (1995). “The Benefits of Mutualism: A Conceptual Framework”. In: *Biological Reviews* 70.3, pp. 427–457.
- Couzin, Iain D, Jens Krause, Richard James, Graeme D Ruxton, and Nigel R Franks (2002). “Collective memory and spatial sorting in animal groups.” In: *Journal of theoretical biology* 218.1, pp. 1–11.
- Cully, Antoine, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret (2015). “Robots that can adapt like animals”. In: *Nature* 521, pp. 503–507.
- D’Ambrosio, David B, Skyler Goodell, Joel Lehman, Sebastian Risi, and Kenneth O Stanley (2012). “Multirobot behavior synchronization through direct neural network communication”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7507 LNAI.PART 2, pp. 603–614.
- D’Ambrosio, David B and Kenneth O Stanley (2008). “Generative Encoding for Multiagent Learning”. In: *Proceedings of the 10th annual conference on Genetic and evolutionary computation GECCO’08* 2008.Gecco, pp. 819–826.

- D'Andrea, Raffaello (2012). "Guest editorial: A revolution in the warehouse: A retrospective on Kiva Systems and the grand challenges ahead". In: *IEEE Transactions on Automation Science and Engineering* 9.4, pp. 638–639.
- Darwin, Charles R (1859). *On the Origin of Species*. London: Murray, John.
- Dawkins, Richard (1976). *The Selfish Gene*. Oxford: Oxford University Press, p. 384.
- DeAngelis, Donald L and Wolf M Mooij (2005). "Individual-Based Modeling of Ecological and Evolutionary Processes". In: *Annual Review of Ecology, Evolution, and Systematics* 36.2005, pp. 147–168.
- Deisenroth, Marc Peter (2011). "A Survey on Policy Search for Robotics". In: *Foundations and Trends in Robotics* 2.1, pp. 1–142.
- Dibangoye, Jilles S., Christopher Amato, Olivier Buffet, and François Charpillet (2015). "Exploiting separability in multiagent planning with continuous-state MDPs". In: *IJCAI International Joint Conference on Artificial Intelligence 2015-Janua.Aamas*, pp. 4254–4260.
- Doncieux, Stephane (2013). "Transfer Learning for Direct Policy Search: A Reward Shaping Approach". In: *Proceedings of the International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. Ieee, pp. 1–6.
- Doncieux, Stephane, Nicolas Bredeche, Jean-Baptiste Mouret, and Agoston E. (Gusz) Eiben (2015). "Evolutionary Robotics: What, Why, and Where to". In: *Frontiers in Robotics and AI* 2.4.
- Doncieux, Stephane and Jean-Baptiste Mouret (2014). "Beyond black-box optimization: a review of selective pressures for evolutionary robotics". In: *Evolutionary Intelligence*, pp. 1–18.
- Dorigo, Marco and Marco Colombetti (1994). "Robot Shaping : Developing Autonomous Agents through Learning". In: *Artificial Intelligence* 70.2, pp. 321–370.
- Drea, Christine M. and Allisa N. Carter (2009). "Cooperative problem solving in a social carnivore". In: *Animal Behaviour* 78.4, pp. 967–977.
- Drea, Christine and Laurence G Frank (2003). "The social complexity of spotted hyenas". In: *Animal Social Complexity: Intelligence, Culture, and Individualized Societies*. Ed. by Frans B M de Waal and Peter T Tyack. Cambridge, Massachusetts: Harvard University Press, pp. 121–148.
- Duarte, Ana, Franz J. Weissing, Ido Pen, and Laurent Keller (2011). "An Evolutionary Perspective on Self-Organized Division of Labor in Social Insects". In: *Annual Review of Ecology, Evolution, and Systematics* 42.1, pp. 91–110.
- Duarte, Miguel, Vasco Costa, Jorge Gomes, Tiago Rodrigues, Fernando Silva, Sancho Moura Oliveira, and Anders Lyhne Christensen (2016). "Evolution of Collective Behaviors for a Real Swarm of Aquatic Surface Robots". In: *PLoS ONE* 11.3.
- Dudek, Gregory, Michael Jenkin, and Evangelos Miliotis (2002). "A Taxonomy of Multirobot Systems". In: *Robot Teams: From Diversity to Polymorphism*. Ed. by Tucker Balch and Lynne E Parker. Vol. 3. 4, pp. 375–397.
- Dugatkin, Lee Alan (1997). *Cooperation Among Animals: An Evolutionary Perspective*. Vol. 53. 9. Oxford University Press.
- (2002). "Cooperation in animals: An evolutionary overview". In: *Biology and Philosophy* 17.4, pp. 459–476.
- Duguid, Shona, Emily Wyman, Anke F Bullinger, Katharina Herfurth-majstorovic, Katharina Herfurth-, and Michael Tomasello (2014). "Coordination strategies of chimpanzees and human children in a Stag Hunt game". In: *Proceedings. Biological sciences / The Royal Society* 281, p. 20141973.

- Eiben, Agoston E. (Gusz) (2013). “Evolutionary robotics”. In: *Communications of the ACM* 56.8, p. 74.
- Eiben, Agoston E and James E Smith (2015). *Introduction to Evolutionary Computing*, p. 299.
- Elena, Santiago F and Richard E Lenski (2003). “Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation.” In: *Nature reviews. Genetics* 4.6, pp. 457–69.
- Eshel, Ilan (1983). “Evolutionary and continuous stability”. In: *Journal of Theoretical Biology* 103.1, pp. 99–111.
- Eshel, Ilan and Uzi Motro (1981). “Kin selection and strong evolutionary stability of mutual help”. In: *Theoretical Population Biology* 19.3, pp. 420–433.
- Eskridge, Brent E., Elizabeth Valle, and Ingo Schlupp (2015). “Emergence of Leadership within a Homogeneous Group”. In: *PloS one* 10.7. Ed. by Cédric Sueur.
- Farinelli, Alessandro, Luca Iocchi, and Daniele Nardi (2004). “Multirobot Systems : A Classification Focused on Coordination”. In: *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* 34.5, pp. 2015–2028.
- Fehr, Ernst and Simon Gächter (2002). “Altruistic punishment in humans.” In: *Nature* 415.6868, pp. 137–40.
- Fernández, Fernando, Daniel Borrajo, and Lynne E Parker (2005). “A reinforcement learning algorithm in cooperative multi-robot domains”. In: *Journal of Intelligent and Robotic Systems: Theory and Applications* 43.2-4, pp. 161–174.
- Ferrante, Eliseo, Ali Emre Turgut, Edgar Duéñez-Guzmán, Marco Dorigo, and Tom Wenseleers (2015). “Evolution of Self-Organized Task Specialization in Robot Swarms.” In: *PLoS computational biology* 11.8.
- Ficici, Sevan G, Richard A Watson, and Jordan B Pollack (1999). “Embodied Evolution: A Response to Challenges in Evolutionary Robotics”. In: *Proceedings of the Eighth European Workshop on Learning Robots* August 2016, pp. 14–22.
- Fisher, R a (1930). “The Genetical Theory of Natural Selection”. In: *Genetics* 154, p. 272.
- Floreano, Dario, Phil Husbands, and Stefano Nolfi (2008). “Evolutionary Robotics”. In: *Springer Handbook of Robotics*. Ed. by Bruno Siciliano and Oussama Khatib. Springer Berlin Heidelberg, pp. 1423–1451.
- Floreano, Dario and Laurent Keller (2010). “Evolution of adaptive behaviour in robots by means of Darwinian selection.” In: *PLoS biology* 8.1, e1000292.
- Floreano, Dario, Sara Mitri, Stéphane Magnenat, and Laurent Keller (2007). “Evolutionary conditions for the emergence of communication in robots.” In: *Current biology : CB* 17.6, pp. 514–519.
- Floreano, Dario and Francesco Mondada (1994). “Automatic Creation of an Autonomous Agent: Genetic Evolution of a Neural-Network Driven Robot”. In: *Proceedings of the Third International Conference on Simulation of Adaptive Behavior : From Animals to Animats 3*. Brighton, United Kingdom: MIT Press, pp. 421–430.
- Floreano, Dario and Stefano Nolfi (1997). “God Save the Red Queen! Competition in Co-Evolutionary Robotics”. In: *Genetic Programming 1997: Proceedings of the Second Annual Conference*, pp. 398–406.
- Floreano, Dario, Stefano Nolfi, and Francesco Mondada (1998). “Competitive Co-Evolutionary Robotics: From Theory to Practice”. In: *From Animals to Animats 5: Proceedings of the 5th International Conference on Simulation of Adaptive Behavior*, pp. 515–524.
- Forber, Patrick and Rory Smead (2015). “Evolution and the classification of social behavior”. In: *Biology & Philosophy*, pp. 1–17.

- Fox, Jeremy W and Richard E Lenski (2015). “From Here to Eternity-The Theory and Practice of a Really Long Experiment.” In: *PLoS biology* 13.6, e1002185.
- Francesca, Gianpiero and Mauro Birattari (2016). “Automatic design of robot swarms: achievements and challenges”. In: *Frontiers in Robotics and AI* 3.29, pp. 1–9.
- Frénoy, Antoine, François Taddei, and Dusan Misevic (2013). “Genetic Architecture Promotes the Evolution and Maintenance of Cooperation”. In: *PLoS Computational Biology* 9.11. Ed. by Claus O. Wilke, e1003339.
- Fukuda, Toshio, Seiya Nakagawa, Yoshio Kawauchi, and Martin Buss (1988). *Self Organizing Robots Based on Cell Structures - CKBOT*.
- Gardner, Andy, Stuart A West, Nicholas H Barton, Andy Gardner, Stuart A West, and Nicholas H Barton (2007). “The Relation between Multilocus Population Genetics and Social Evolution Theory”. In: *The American Naturalist* 169.2, pp. 207–226.
- Geritz, Stefan A H, Éva Kisdi, Geza Meszèna, and Johan Metz (1998). “Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree”. In: Geritz, Stefan a H, Éva Kisdi, Géza Meszéna, and Johan a J Metz (2004). “Adaptive Dynamics of Speciation: Ecological Underpinnings”. In: *Adaptive Speciation*. Ed. by Ulf Dieckmann, M Doebeli, Johan a J Metz, and D Tautz. November. Cambridge University Press, pp. 54–75.
- Goldberg, David E (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*.
- Goldsby, Heather J, Anna Dornhaus, Benjamin Kerr, and Charles Ofria (2012). “Task-switching costs promote the evolution of division of labor and shifts in individuality.” In: *Proceedings of the National Academy of Sciences of the United States of America* 109.34, pp. 13686–13691.
- Goldsby, Heather J, David B Knoester, and Charles Ofria (2010). “Evolution of Division of Labor in Genetically Homogenous Groups”. In: *Proceedings of the 12th annual conference on Genetic and evolutionary computation - GECCO '10*. ACM, pp. 135–142.
- Grafen, Alan (1984). *Natural Selection, Kin Selection and Group Selection*. Ed. by John R. Krebs and Nicholas B Davies. Oxford, UK: Blackwell Scientific Publications, pp. 62–84.
- Grimm, V and Sf Railsback (2005). “Individual-based Modeling and Ecology”. In: *Individual-based Modeling and Ecology*, pp. 1–432.
- Groß, Roderich and Marco Dorigo (2004). “Cooperative Transport of Objects of Different Shapes and Sizes”. In: *Ant Colony Optimization and Swarm Intelligence: 4th International Workshop, ANTS 2004, Brussels, Belgium, September 5-8, 2004. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 106–117.
- Gueron, Shay, Simon A Levin, and Daniel L Rubenstein (1996). “The dynamics of herds: From individuals to aggregations”. In: *Journal of Theoretical Biology* 182.1, pp. 85–98.
- Haley, Patrick B, Randal S Olson, Fred C Dyer, and Christoph Adami (2014). “Exploring conditions that select for the evolution of cooperative group foraging”. In: *Proceedings of the 14th International Conference on the Synthesis and Simulation of Living Systems (ALIFE 14)*, pp. 310–311.
- Hamilton, William D (1964). “The genetical evolution of social behaviour. I.” In: *Journal of theoretical biology* 7.1, pp. 1–16.
- (1971). “Geometry for the selfish herd”. In: *Journal of Theoretical Biology* 31.2, pp. 295–311.
- (1975). “Innate social aptitudes of man: an approach from evolutionary genetics”. In: *Biosocial Anthropology*. Ed. by Robin Fox. London: Malaby Press, pp. 133–153.
- Hammerstein, Peter and Reinhard Selten (1994). “Game theory and evolutionary biology”. In: *Handbook of Game Theory with Economic Applications* 2, pp. 929–993.
- Hardin, Garrett (1968). “The Tragedy of the Commons”. In: *Science* 162.3859, pp. 1243–1248.

- Harrison, Freya (2013). “Bacterial cooperation in the wild and in the clinic: Are pathogen social behaviours relevant outside the laboratory?” In: *BioEssays* 35.2, pp. 108–112.
- Hart, Benjamin L and Lynette A Hart (1992). “Reciprocal allogrooming in impala, *Aepyceros melampus*”. In: *Animal Behaviour* 44.6, pp. 1073–1083.
- Harvey, Inman, Phil Husbands, and Dave Cliff (1994). “Seeing The Light : Artificial Evolution, Real Vision”. In: *From Animals to Animats 3: Proc. 3rd Int. Conf. on Simulation of Adaptive Behavior*. Vol. 1994. April, pp. 392–401.
- Hauert, Christoph and Michael Doebeli (2004). “Spatial structure often inhibits the evolution of cooperation in the snowdrift game”. In: *Nature* 428.6983, pp. 643–646.
- Hauert, Sabine, Severin Leven, Maja Varga, Fabio Ruini, Angelo Cangelosi, Jean Christophe Zufferey, and Dario Floreano (2011). “Reynolds flocking in reality with fixed-wing robots: Communication range vs. maximum turning rate”. In: *IEEE International Conference on Intelligent Robots and Systems*, pp. 5015–5020.
- Hauert, Sabine, Sara Mitri, Laurent Keller, and Dario Floreano (2014). “Evolving Cooperation : From Biology to Engineering”. In: *The Horizons of Evolutionary Robotics*. Ed. by P Vargas, E Di parolo, Inman Harvey, and Phil Husbands. Cambridge, MA, USA: MIT Press, pp. 203–217.
- Hauert, Sabine, Jean Christophe Zufferey, and Dario Floreano (2009). “Evolved swarming without positioning information: An application in aerial communication relay”. In: *Autonomous Robots* 26.1, pp. 21–32.
- Hertzberg, Joachim and Raja Chatila (2008). “AI Reasoning Methods for Robotics”. In: *Springer Handbook of Robotics*, pp. 207–223.
- Holland, John H (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*.
- Huston, Michael, Donald DeAngelis, and Wilfred Post (1988). “New Computer Models Unify Ecological Theory”. In: *BioScience* 38.10, pp. 682–691.
- Huth, Andreas and Christian Wissel (1992). “The Simulation of the Movement of Fish Schools”. In: *Journal of Theoretical Biology* 156.3, pp. 365–385.
- Iocchi, Luca, Daniele Nardi, and Massimiliano Salerno (2001). “Reactivity and Deliberation: A Survey on Multi-Robot Systems”. In: *Balancing Reactivity and Social Deliberation in Multi-Agent Systems, From RoboCup to Real-World Applications (Selected Papers from the ECAI 2000 Workshop and Additional Contributions)*, pp. 9–34.
- Iyer, Swami and Timothy Killingback (2016). “Evolution of Cooperation in Social Dilemmas on Complex Networks”. In: *PLoS computational biology* 12.2.
- Jakobi, Nick, Phil Husbands, and Inman Harvey (1995). “Noise and the reality gap: The use of simulation in evolutionary robotics”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 929, pp. 704–720.
- Keller, Laurent and Michael G Surette (2006). “Communication in bacteria: an ecological and evolutionary perspective”. In: *Nature reviews Microbiology* 4, pp. 249–258.
- Kitano, Hiroaki, Minoru Asada, Yasuo Kuniyoshi, Itsuki Noda, Eiichi Osawa, and Hitoshi Matsubara (1997). “RoboCup: A challenge problem for AI”. In: *AI Magazine* 18.1, pp. 1–19.
- Knibbe, C, G Beslon, V Lefort, F Chaudier, and J Fayard (2005). “Self-adaptation of Genome Size in Artificial Organisms”. In: pp. 423–432.
- Kober, Jens, J Andrew Bagnell, and Jan Peters (2013). “Reinforcement learning in robotics: A survey”. In: *The International Journal of Robotics Research* 32.11, pp. 1238–1274.

- Kube, C. Ronald and Hong Zhang (1993). “Collective Robotics: From Social Insects to Robots”. In: *Adaptive Behavior* 2.2, pp. 189–218.
- Lehman, Joel and Kenneth O Stanley (2008). “Exploiting Open-Endedness to Solve Problems Through the Search for Novelty”. In: *Artificial Life XI* Alife Xi, pp. 329–336.
- (2011). “Abandoning Objectives: Evolution through the Search for Novelty Alone”. In: *Evolutionary computation* 19.2, pp. 189–223.
- (2013). “Evolvability is inevitable: increasing evolvability without the pressure to adapt.” In: *PloS one* 8.4, e62186.
- Lehmann, Laurent and Laurent Keller (2006). “The evolution of cooperation and altruism - A general framework and a classification of models”. In: *Journal of Evolutionary Biology* 19.5, pp. 1365–1376.
- Leimar, Olof and Richard C Connor (2003). “By-product Benefits, Reciprocity, and Pseudoreciprocity in Mutualism”. In: *Genetic and Cultural Evolution of Cooperation*. Ed. by P. Hammerstein. MIT Press, pp. 203–222.
- Leimar, Olof and Peter Hammerstein (2010). “Cooperation for direct fitness benefits.” In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 365.1553, pp. 2619–26.
- Lenski, Richard E, Charles Ofria, Travis C Collier, and Christoph Adami (1999). “Genome complexity, robustness and genetic interactions in digital organisms”. In: *Nature* 400.6745, pp. 661–664.
- Lichocki, Pawel, Steffen Wischmann, Laurent Keller, and Dario Floreano (2013). “Evolving team compositions by agent swapping”. In: *IEEE Transactions on Evolutionary Computation* 17.2, pp. 282–298.
- Littman, Michael L (1994). “Markov games as a framework for multi-agent reinforcement learning”. In: *Proceedings of the International Conference on Machine Learning* 157.1, pp. 157–163.
- Matarić, Maja J (1995). “Issues and Approaches in the Design of Collective Autonomous Agents”. In: *Robotics and Autonomous Systems* 16.2-4, pp. 321–331.
- (1997). “Reinforcement Learning in the Multi-Robot Domain”. In: *Autonomous Robots* 4, pp. 73–83.
- Matarić, Maja J and François Michaud (2008). “Behavior-Based Systems”. In: *Springer Handbook of Robotics*, pp. 891–909.
- Maynard Smith, John (1976). “Group Selection”. In: *Quarterly Review of Biology* 51, pp. 277–283.
- (1982). *Evolution and the Theory of Games*. London: Cambridge University Press.
- Maynard Smith, John and G R Price (1973). “Logic of Animal Conflict”. In: *Nature* 246.5427, pp. 15–18.
- McGill, Brian J. and Joel S. Brown (2007). “Evolutionary Game Theory and Adaptive Dynamics of Continuous Traits”. In: *Annual Review of Ecology, Evolution, and Systematics* 38.1, pp. 403–435.
- Mills, M G L (2003). *Kalahari Hyenas Comparative Behavioral Ecology of Two Species*. The Blackburn Press, p. 320.
- Mitri, Sara, Dario Floreano, and Laurent Keller (2009). “The evolution of information suppression in communicating robots with conflicting interests.” In: *Proceedings of the National Academy of Sciences* 106.37, pp. 15786–15790.
- (2011). “Relatedness influences signal reliability in evolving robots”. In: *Proceedings of the Royal Society B: Biological Sciences* 278.1704, pp. 378–383.

- Mitri, Sara, Steffen Wischmann, Dario Floreano, and Laurent Keller (2013). “Using robots to understand social behaviour.” In: *Biological reviews of the Cambridge Philosophical Society* 88.1, pp. 31–9.
- Mondada, Francesco, Luca Maria Gambardella, Dario Floreano, Stefano Nolfi, Jean Louis Deneubourg, and Marco Dorigo (2005). “The cooperation of swarm-bots: Physical interactions in collective robotics”. In: *IEEE Robotics and Automation Magazine* 12.2, pp. 21–28.
- Montague, P R, P Dayan, and T J Sejnowski (1996). “A framework for mesencephalic dopamine systems based on predictive Hebbian learning.” In: *Journal of neuroscience* 16.5, pp. 1936–1947.
- Montanier, Jean-Marc and Nicolas Bredeche (2011). “Surviving the Tragedy of Commons : Emergence of Altruism in a Population of Evolving Autonomous Agents”. In: *Proceedings of the 11th International Conference on Artificial Life (ECAL'11)*, pp. 550–557.
- (2013). “Evolution of Altruism and Spatial Dispersion: an Artificial Evolutionary Ecology Approach”. In: *Advances in Artificial Life, ECAL 2013*, pp. 260–267.
- Mouret, Jean-Baptiste and Stephane Doncieux (2010). “SFERESv2: Evolvin’ in the Multi-Core World”. In: *Proceedings of Congress on Evolutionary Computation (CEC)*. 2. Ieee, pp. 4079–4086.
- (2012). “Encouraging Behavioral Diversity in Evolutionary Robotics: An Empirical Study”. In: *Evolutionary Computation* 20.1, pp. 91–133.
- Mouret, Jean-Baptiste, Sylvain Koos, and Stephane Doncieux (2012). “Crossing the Reality Gap: a Short Introduction to the Transferability Approach”. In: *Proceedings of the first workshop on the Evolution of Physical Systems (EPS)*, pp. 1–7.
- Mouret, Jean-Baptiste and Paul Tonelli (2014). “Artificial Evolution of Plastic Neural Networks: a few Key Concepts”. In: *Growing Adaptive Machines: combining Development and Learning in Artificial Neural Networks*, pp. 251–261.
- Nash, John F (1950). “Equilibrium points in n-person games”. In: *Proceedings of the National Academy of Sciences* 36.1, pp. 48–49.
- Nilsson, Nils J (1984). *Shakey the Robot*. Tech. rep.
- Nitschke, G. S., M. C. Schut, and A. E. Eiben (2012). “Evolving behavioral specialization in robot teams to solve a collective construction task”. In: *Swarm and Evolutionary Computation* 2.January, pp. 25–38.
- Nolfi, Stefano and Dario Floreano (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press.
- Nunney, Leonard (1989). “The Maintenance of Sex by Group Selection”. In: *Evolution* 43.2, pp. 245–257.
- Olson, Randal S, Patrick B Haley, Fred C Dyer, Christoph Adami, and East Lansing (2014). “Exploring the evolution of a trade-off between vigilance and foraging in group-living organisms”. In: August, pp. 1–26.
- Olson, Randal S, Arend Hintze, Fred C Dyer, David B Knoester, and Christoph Adami (2013a). “Predator confusion is sufficient to evolve swarming behavior”. In: *Journal of the Royal Society Interface* 10.85, p. 20130305.
- Olson, Randal S, Arend Hintze, Fred C Dyer, Jason H Moore, and Christoph Adami (2016). “Exploring the coevolution of predator and prey morphology and behavior”. In:
- Olson, Randal S, David B Knoester, and Christoph Adami (2013b). “Critical interplay between density-dependent predation and evolution of the selfish herd”. In: *Proceeding of the fifteenth annual conference on Genetic and evolutionary computation conference - GECCO '13*, p. 247.

- Panait, Liviu and Sean Luke (2005). “Cooperative Multi-Agent Learning: The State of the Art”. In: *Autonomous Agents and Multi-Agent Systems* 11, pp. 387–434.
- Parker, Lynne E (1994). “Heterogeneous Multi-Robot Cooperation”. In: *Jet Propulsion* February, p. 227.
- (2000). “Current state of the art in distributed autonomous mobile robotics”. In: *Distributed autonomous robotic systems*, pp. 3–12.
- (2008). “Multiple mobile robot systems”. In: *Springer Handbook of Robotics*. Ed. by Bruno Siciliano and Oussama Khatib. Springer Berlin Heidelberg, pp. 921–941.
- Perkins, Simon and Gillian Hayes (1996). *Robot Shaping - Principles, Methods and Architectures*. Tech. rep. Edimbourg, GB: University of Edinburgh.
- Popat, Roman, Shanika A Crusz, Marco Messina, Paul Williams, Stuart A West, and Stephen P Diggle (2012). “Quorum-sensing and cheating in bacterial biofilms.” In: *Proceedings. Biological sciences / The Royal Society* 279.1748, pp. 4765–71.
- Potter, Mitchell A and Kenneth A De Jong (1994). “A cooperative coevolutionary approach to function optimization”. In: *Parallel Problem Solving from Nature - PPSN III*, pp. 249–257.
- Poulin, Robert and Alexandra S Grutter (1996). “Cleaning Symbioses: Proximate and Adaptive Explanations”. In: *BioScience* 46.7, pp. 512–517.
- Queller, Dc and Je Strassmann (1998). “Kin selection and social insects”. In: *BioScience* 48.8, pp. 165–175.
- Quinn, M. (2001). “A comparison of approaches to the evolution of homogeneous multi-robot teams”. In: *Proceedings of the 2001 Congress on Evolutionary Computation*, pp. 128–135.
- Requejo-Martinez, Rubén J (2013). “Evolutionary game theory and the tower of Babel of cooperation: Altruism, free-riding, parasitism and the structure of the interactions in a world with finite resources.” In:
- Reynolds, C. W. (1992). “An Evolved, Vision-Based Behavioral Model of Coordinated Group Motion”. In: *Proceedings of the Second International Conference on Simulation of Adaptive Behaviour (SAB92)*, pp. 384–392.
- Reynolds, Craig W. (1987). “Flocks, herds and schools: A distributed behavioral model”. In: *ACM SIGGRAPH Computer Graphics* 21.4, pp. 25–34.
- Rubenstein, Michael, Alejandro Cornejo, and Radhika Nagpal (2014). “Programmable self-assembly in a thousand-robot swarm”. In: *Science* 345.6198, pp. 795–799.
- Sachs, Joel L, Ulrich G. Mueller, Thomas P. Wilcox, and James J. Bull (2004). “The evolution of cooperation”. In: *The Quarterly review of biology* 51.2, pp. 211–44.
- Saksida, Lisa M, Scott M Raymond, and David S Touretzky (1997). “Shaping robot behavior using principles from instrumental conditioning”. In: *Robotics and Autonomous Systems* 22, pp. 231–249.
- Scott-Phillips, Thomas C, Thomas E Dickins, and Stuart A West (2011). “Evolutionary Theory and the Ultimate-Proximate Distinction in the Human Behavioral Sciences”. In: *Perspectives on Psychological Science* 6.1, pp. 38–47.
- Sen, Sandip, Mahendra Sekaran, and John Hale (1994). “Learning to coordinate without sharing information”. In: *Proceedings of the National Conference on Artificial Intelligence*, pp. 426–431.
- Servedio, Maria R., Yaniv Brandvain, Sumit Dhole, Courtney L. Fitzpatrick, Emma E. Goldberg, Caitlin A. Stern, Jeremy Van Cleve, and D. Justin Yeh (2014). “Not Just a Theory - The Utility of Mathematical Models in Evolutionary Biology”. In: *PLoS Biology* 12.12, pp. 1–5.

- Seuken, Sven and Shlomo Zilberstein (2007). “Memory-bounded dynamic programming for DEC-POMDPs”. In: *IJCAI International Joint Conference on Artificial Intelligence*, pp. 2009–2015.
- Shou, Wenying, Carl T Bergstrom, Arup K Chakraborty, and Frances K Skinner (2015). “Theory, models and biology”. In: *eLife* 4.
- Skyrms, Brian (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press.
- Smaldino, Paul E (2014). “The cultural evolution of emergent group-level traits”. In: *Behav Brain Sci* 37.3, pp. 243–254.
- Smith, Jennifer E., Eli M. Swanson, Daphna Reed, and Kay E. Holekamp (2012). “Evolution of Cooperation among Mammalian Carnivores and Its Relevance to Hominin Evolution”. In: *Current Anthropology* 53.S6, S436–S452.
- Smith, Jennifer E., Russell C. Van Horn, Katherine S. Powning, Alison R. Cole, Katharine E. Graham, Sandra K. Memenis, and Kay E. Holekamp (2010). “Evolutionary forces favoring intragroup coalitions among spotted hyenas and other animals”. In: *Behavioral Ecology* 21.2, pp. 284–303.
- Smith, Mark (1991). “Using Massively-Parallel Supercomputers to Model Stochastic Spatial Predator-Prey Systems”. In: *International Journal of Ecological Modelling*.
- Solomon, Michael, Terence Soule, and Robert B. Heckendorn (2012). “A comparison of a communication strategies in cooperative learning”. In: *Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference - GECCO '12*, p. 153.
- Stanley, Kenneth O, David D Ambrosio, and Jason Gauci (2009). “A Hypercube-Based Indirect Encoding for Evolving Large-Scale Neural Networks”. In: *Artificial Life journal* 15.2, pp. 1–39.
- Stanley, Kenneth O and Risto Miikkulainen (2002). “Evolving Neural Networks through Augmenting Topologies”. In: *Evolutionary Computation* 10.2, pp. 99–127.
- Stone, Peter and Rs Sutton (2001). “Scaling reinforcement learning toward RoboCup soccer”. In: *Icml* June, pp. 537–544.
- Stone, Peter and Manuela Veloso (2000). “Multiagent Systems : A Survey from a Machine Learning Perspective”. In: *Autonomous Robotics* 8.3, pp. 345–383.
- Stulp, Freek and Olivier Sigaud (2013). “Robot Skill Learning: From Reinforcement Learning to Evolution Strategies”. In: *Paladyn, Journal of Behavioral Robotics* 4.1, pp. 49–61.
- Sutton, R S and a G Barto (1998). *Reinforcement Learning: An Introduction*. Vol. 9, pp. 1054–1054.
- Sutton, Richard S (1988). “Learning to Predict by the Method of Temporal Differences”. In: *Machine Learning* 3.1, pp. 9–44.
- Sutton, Richard S., David Mcallester, Satinder Singh, and Yishay Mansour (2000). “Policy Gradient Methods for Reinforcement Learning with Function Approximation”. In: *Advances in Neural Information Processing Systems* 12. Vol. 12, pp. 1057–1063.
- Szathmàry, Eörs and John Maynard Smith (1995). “The major evolutionary transistions”. In: *Nature* 374.6519, pp. 227–232.
- Tarapore, Danesh, Dario Floreano, and Laurent Keller (2010). “Task-dependent influence of genetic architecture and mating frequency on division of labour in social insect societies”. In: *Behavioral Ecology and Sociobiology* 64.4, pp. 675–684.
- Taylor, Tim, Mark Bedau, Alastair Channon, David Ackley, Wolfgang Banzhaf, Guillaume Beslon, Emily Dolson, Tom Froese, Simon Hickinbotham, Takashi Ikegami, Barry McMullin,

- Norman Packard, Steen Rasmussen, Nathaniel Virgo, Eran Agmon, Edward Clark, Simon McGregor, Charles Ofra, Glen Ropella, Lee Spector, Kenneth O Stanley, Adam Stanton, Christopher Timperley, Anya Vostinar, and Michael Wiser (2016). “Open-Ended Evolution: Perspectives from the OEE Workshop in York”. In: *Artificial Life* 22.3, pp. 408–423.
- Tinbergen, Niko (1963). *On aims and methods of Ethology*.
- Tomasello, Michael, Alicia P. Melis, Claudio Tennie, Emily Wyman, and Esther Herrmann (2012). “Two Key Steps in the Evolution of Human Cooperation.” In: *Current Anthropology* 53.6, pp. 673–692.
- Toquenaga, Yukihiro, Isamu Kajitani, and Tsutomu Hoshino (1995). “Egrets of a Feather Flock Together”. In: *Artificial Life*. Ed. by Mark A Bedau. Vol. 1. 4, pp. 391–411.
- Trianni, Vito (2014). “Evolutionary Robotics: Model or Design?” In: *Frontiers in Robotics and AI* 1.December, pp. 1–6.
- Trianni, Vito, Christos Ampatzis, Anders Lyhne Christensen, Elio Tuci, Marco Dorigo, and Stefano Nolfi (2007). “From Solitary to Collective Behaviours : Decision Making and Cooperation”. In: *Advances in Artificial Life*, pp. 575–584.
- Trianni, Vito, Roderich Groß, Thomas H Labella, Erol Şahin, and Marco Dorigo (2003). “Evolving Aggregation Behaviors in a Swarm of Robots”. In: *ECAL 2003*, pp. 865–874.
- Trianni, Vito, Stefano Nolfi, and Marco Dorigo (2004). *Cooperative Hole Avoidance in a Swarm-bot*. Tech. rep. October, pp. 1–12.
- Trivers, Robert (1971). “The Evolution of Reciprocal Altruism”. In: *The Quarterly Review of Biology* 46.1, pp. 35–57.
- (1985). *Social Evolution*. Benjamin-Cummings.
- Tuci, Elio and Vito Trianni (2014). “On the evolution of homogeneous two-robot teams: clonal versus aclonal approaches”. In: *Neural Computing and Applications* 25.5, pp. 1063–1076.
- Urzelai, Joseba and Dario Floreano (2001). “Evolution of adaptive synapses: robots with fast adaptive behavior in new environments.” In: *Evolutionary computation* 9.4, pp. 495–524.
- Urzelai, Joseba, Dario Floreano, Marco Dorigo, and Marco Colombetti (1998). “Incremental Robot Shaping”. In: *Connection Science* 10.3-4, pp. 341–360.
- Van Baalen, Minus and David A Rand (1998). “The unit of selection in viscous populations and the evolution of altruism”. In: *Journal of Theoretical Biology* 193.4, pp. 631–648.
- Vienne, Damien M. de, Tatiana Giraud, and Pierre Henri Gouyon (2013). “Lineage Selection and the Maintenance of Sex”. In: *PLoS ONE* 8.6.
- Wagner, Gunter P and Lee Altenberg (1996). “Perspective: Complex Adaptations and the Evolution of Evolvability”. In: *Evolution* 50.3, pp. 967–976.
- Waibel, Markus, Dario Floreano, and Laurent Keller (2011). “A quantitative test of Hamilton’s rule for the evolution of altruism.” In: *PLoS biology* 9.5, e1000615.
- Waibel, Markus, Dario Floreano, Stéphane Magnenat, and Laurent Keller (2006). “Division of labour and colony efficiency in social insects: effects of interactions between genetic architecture, colony kin structure and rate of perturbations.” In: *Proceedings. Biological sciences / The Royal Society* 273.1595, pp. 1815–23.
- Waibel, Markus, Laurent Keller, Dario Floreano, and Senior Member (2009). “Genetic Team Composition and Level of Selection in the Evolution of Cooperation”. In: *IEEE Transactions on Evolutionary Computation* 13.3, pp. 648–660.
- Watkins, Christopher J C H (1989). *Chris Watkins’ PhD thesis*.

- Watson, Richard A, Sevan G Ficici, and Jordan B Pollack (2002). “Embodied Evolution: Distributing an evolutionary algorithm in a population of robots”. In: *Robotics and Autonomous Systems* 39.1, pp. 1–18.
- Watson, Richard A and Jordan B Pollack (2000). “Mutualism , Parasitism , and Evolutionary Adaptation”. In:
- Werfel, Justin, Kirstin Petersen, and Radhika Nagpal (2014). “Designing collective behavior in a termite-inspired robot construction team.” In: *Science (New York, N.Y.)* 343.6172, pp. 754–8.
- Werger, Barry B and Maja J Matarić (2000). “Broadcast of Local Eligibility for Multi-Target Observation”. In: *5th International Conference on Distributed Autonomous Robotic Systems DARS00 4*, pp. 347–356.
- West, Stuart A, Claire El Mouden, and Andy Gardner (2011). “Sixteen common misconceptions about the evolution of cooperation in humans”. In: *Evolution and Human Behavior* 32.4, pp. 231–262.
- West, Stuart A, Ashleigh S Griffin, and Andy Gardner (2007a). “Evolutionary explanations for cooperation.” In: *Current biology : CB* 17.16, R661–72.
- (2007b). “Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection.” In: *Journal of evolutionary biology* 20.2, pp. 415–32.
- West, Stuart A, Ashleigh S Griffin, Andy Gardner, and Stephen P Diggle (2006). “Social evolution theory for microorganisms”. In: *Nat.Rev.Microbiol.* 4.1740-1526 (Print), pp. 597–607.
- Whiteson, Shimon (2012). “Evolutionary Computation for Reinforcement Learning”. In: *Reinforcement Learning: State of the Art* 12.c, pp. 325–355.
- Wilder, Bryan and Kenneth O Stanley (2015). “Altruists Proliferate Even at a Selective Disadvantage within Their Own Niche”. In: *Plos One* 10.6, e0128654.
- Wilson, Edward O (1990). *Success and Dominance in Ecosystems: The Case of the Social Insects*. Oldendorf/Luhe, Federal Republic of Germany: Ecology Institute.
- Wilson, W G (1998). “Resolving discrepancies between deterministic population models and individual-based simulations.” In: *The American naturalist* 151.2, pp. 116–134.
- Wischmann, Steffen, Dario Floreano, and Laurent Keller (2012). “Historical contingency affects signaling strategies and competitive abilities in evolving populations of simulated robots”. In: *Proceedings of the National Academy of Sciences* 109.3, pp. 864–868.
- Wrangham, R W, J H Jones, G Laden, D Pilbeam, and N Conklin-Brittain (1999). “The Raw and the Stolen. Cooking and the Ecology of Human Origins.” In: *Current Anthropology* 40.5, pp. 567–594.
- Wright, Jonathan, Elena Berg, Selvino R D E Kort, Vladimir Khazin, and Alexei a Maklakov (2001). “Cooperative sentinel behaviour in the Arabian babbler”. In: *Animal Behaviour* 61, pp. 1–7.
- Wright, Sewall (1931). “Evolution in Mendelian Populations”. In: *Genetics* 16.2, pp. 97–159.
- Yang, Erfu and Dougbing Gu (2005). “A Survey on Multiagent Reinforcement Learning Towards Multi-Robot Systems.” In: *Proceedings of the IEEE symposium on Computational Intelligence and Games (CIG) 2005 2*.
- Yong, Chern Han and Risto Miikkulainen (2009). “Coevolution of role-based cooperation in multi-agent systems”. In: *IEEE Transactions on Autonomous Mental Development* 1.3, pp. 170–186.

