Tel Aviv University
Raymond and Beverly Sackler Faculty of Exact Sciences
School of Computer Science

Evolutionary Dynamics of Adaptive Populations:
The Effect of Phenotypic Plasticity, Imitation and Culture on Evolution

by

Elhanan Borenstein

under the supervision of Prof. Eytan Ruppin and Prof. Marcus Feldman

A thesis submitted
for the degree of Doctor of Philosophy

Submitted to the Senate of Tel Aviv University
September 2006

Acknowledgements

This dissertation summarizes a wonderful period I spent at Tel-Aviv University. I was fortunate to study and work with many gifted people who enriched me in numerous and often unexpected ways. First and foremost, I am greatly indebted to Eytan Ruppin and Marcus Feldman, my supervisors. I feel extremely privileged to have worked with such brilliant and good-hearted scientists. My greatest debt is to Eytan, for taking me under his wing from my very first steps in the academia, teaching me what science and research are all about, and being a patient and generous mentor. This debt could never be paid in full, and I can only hope to repay some of it when we meet further down the road. In an introduction to one of his talks, someone once presented Eytan as the ideal classical scientist - one that follows his heart in pursuit of fresh and exciting research questions. If I was forced to name just one trait I have learned from working with Eytan, that would be the one. It was a pleasure to work with a supervisor that granted me the freedom to pursue the topics I am passionate about, and offered his full support even when those were not fully aligned with his own research interests. I am also grateful to Eytan for teaching me good scientific judgment and the relentlessness with which scientific research should be conducted and presented. Above all, I am thankful for his genuine friendship. Marc was kind enough to join Eytan in supervising my PhD and generously invited me to spend some time at Stanford, based initially on not much more than a short phone conversation we had. I hope that in the time that passed since that conversation took place I was able to prove I was worthy of his confidence. Marc guided me in the world of formal population biology, taught me the importance of meticulous analysis, and helped me make the transition from a computer scientist working on biologically motivated questions into a theoretical biologist using computer science as a tool of research. Even from a distance, he was able to offer valuable suggestions and support throughout my PhD. Eytan's and Marc's guidance has made me the sci-

entist I am today, and they will always be my model for the kind of scientist I hope to become.

I also owe a special debt to Isaco Meilijson for taking the time and effort to discuss with me several statistical and mathematical issues. It gave me a rare peek into the amazing mind and thinking of a true mathematician, and taught me that intuition is just as valuable a tool as formalism in tackling mathematical problems. One such discussion that lasted over a period of several months has lead the way to our joint publication.

During my PhD studies, I have enjoyed the opportunity of spending a few months in other academic institutes. I am thankful to the Santa Fe Institute and the Philip Steinmetz Fellowship for allowing me to attend the Complex Systems Summer School and to stay at the Santa-Fe Institute for a couple of months. These experiences have introduced to me the joy and excitement of a truly multidisciplinary environment unlike any other I have ever been in. I owe a debt of gratitude to Prof. David Krakauer for taking me under his wing during my visit and for allowing (and often pushing) me to think outside the box. I envy his extremely diverse knowledge, his endless curiosity, and mostly, his overwhelming creativity, as well as his remarkable skills in Halo. Conversations over lunch with Jessica Flack, Eric Smith, Jon Wilkins and many other members of the Institute were always mind opening. I thank the members of Feldman's lab at Stanford for their hospitality and warm welcome during the summers I spent at Stanford University: Jeremy Kendal (who was also my partner in one of the papers included in this dissertation), Lilach Hadany, Aaron Hirsh and Joanna Masel.

Being a part of a cheerful and talented group of people is a joyful experience, and helps to overcome the occasional loneliness associated with scientific work. I would like to thank the members of Eytan's lab, Zach Solan, Alon Kaufmann, Anat Kreimer, Vered Kunik, Yossi Mossel, Liron Levkovitz and Jonathan Berant for being there for me whenever I needed them. I especially want to thank my office roommates, Alon Keinan, Tomer Shlomi and Ben Sandbank, with whom I have had numerous stimulating discussions (though, most of which, had nothing to do with science), for making our office a noisy but mostly amusing place to work in. Ranit Aharonov, a former member of Eytan's lab, guided me through my very first steps in the lab, and many years later, during a long car ride from Jerusalem to Tel-Aviv, helped initiate my work on MicroRNA robustness (not included in this dissertation) .

One of the unexpected delights of the academic life is the occasional encounters with gifted and exceptional people. I thank Arnon Lotem for several stimulating discussions. Gal Chechik, Daniel Polani (who also read several early versions of my manuscripts), Guy Sella, Ronny Meir, David Horn, Sid Strauss, Omer Berkman and Fyodor Kondrashov were each kind enough to share with me their views and gave me a unique perspective on scientific issues.

The research described in this thesis was supported by several external funding sources. The Yeshaya Horowitz Association through the Center of Complexity Science provided me with a full support scholarship for the last two years of my PhD. Aharon Katzir Center Training Fellowships and the Don & Sara Marejn Scholarship provided additional generous support.

A special thank you is due to Tova Kalish and Pnina Barzilay, who were there for me whenever an administrative issue came up, and saved me the agonies of academic bureaucracy.

Finally, I wish to thank my friends and family for their ongoing support, love and friendship.

In memory of my mother,
Shoshana Borenstein,
the cause of everything I am,
and everything I ever will be.

Abstract

When facing the challenge of developing an individual that best fits its environment, nature demonstrates an interesting combination of two, fundamentally different, adaptive mechanisms: Genetic evolution and lifetime learning. Although lifetime adaptations are not inherited in a pure Darwinian framework, they may still change the individual's fitness, and consequently the expected number of offspring that carry its genotype. Such lifetime learning mechanisms thus dramatically alter the course of the evolutionary process. When learning is implemented via imitative or cultural learning, the resulting dynamics are further complicated as the success of the learning process depends not only on the interaction between the individual and its environment but also on the state of other members of the evolving population. The primary goal of this dissertation is hence to examine the complex interplay between evolution, learning, imitation and culture.

In the first study presented in this dissertation, we examine the effect of simple phenotypic plasticity on the rate of genetic evolution. Following numerous computational models, it has become the accepted wisdom that lifetime acclimation (e.g., via learning) smooths the fitness landscape and consequently accelerates evolution. However, analytical studies, focusing on the effect of phenotypic plasticity on evolution in simple unimodal landscapes, have often found that learning hinders the evolutionary process rather than accelerating it. Addressing this discrepancy, we provide a general framework for studying the effect of plasticity on evolution in multipeaked landscapes and introduce a rigorous mathematical analysis of these dynamics. We show that the convergence rate of the evolutionary process in a given arbitrary one-dimensional fitness landscape is dominated by the largest descent (drawdown) in the landscape and provide numerical evidence to support an analogous dominance also in multidimensional landscapes. We consider several schemes of phenotypic plasticity and individual learning and examine their effect on the landscape drawdown, identifying the conditions under which plasticity is advantageous. The lack of such a drawdown in unimodal landscapes vs. its dominance in multipeaked landscapes accounts for the seemingly

contradictory findings of previous studies.

The plasticity schemes and evolutionary model examined in the first study facilitate a rigorous quantitative analysis of the effect of plasticity on evolution. Such an analysis is limited to simple learning models wherein the effect of learning on the fitness landscape structure can be clearly characterized. However, living organisms demonstrate a variety of complex learning mechanisms, posing a greater modeling and analysis challenge. Specifically, learning by imitation is a highly complex cognitive process, involving vision, perception, representation, memory and motor control and has attracted a great deal of attention in recent years. In the second and third studies, we thus focus on the interaction between learning by imitation and evolution, using the framework of Evolutionary Autonomous Agents (EAA).

We first focus on the effect of learning by imitation on the evolutionary process. We describe a set of simulations where a population of agents evolves to solve a certain task. In each generation, individuals can select other agents from the population as models (teachers) and imitate their behavior. In contradistinction to previous studies, we focus on the interaction between imitation and evolution when imitation takes place only across members of the same generation, and does not percolate across generations via vertical (cultural) transmission. We show that introducing such an imitative process successfully enhances the evolution of autonomous agents when other forms of learning are not applicable.

Acknowledging the bidirectional interplay between evolution and imitation, we then turn to examine the evolution of the underlying mechanisms that give rise to imitative behavior. These mechanisms have been the subject of research in various disciplines, from neuroscience to animal behavior and human psychology. In particular, studies in monkeys and humans have discovered a neural mirror system that demonstrates an internal correlation between the representations of perceptual and motor functionalities. In contrast to previous engineering-based approaches, we present a novel framework for studying the evolutionary emergence of imitative behavior. We develop evolutionary adaptive agents that demonstrate imitative learning, facilitating a comprehensive study of the emerging underlying neural mechanisms. Interestingly, these agents are found to include a neural

"mirror" device comparable to those identified in biological systems. Further analysis of these agents' networks reveals complex dynamics, combining innate perceptual-motor coupling with acquired context-action associations, to accomplish the required task. These findings suggest a universal and fundamental link between the ability to replicate the actions of others (imitation) and the capacity to represent and match others' actions (mirroring).

Imitative learning capabilities may eventually give rise to cultural evolution - the evolution of ideas, thoughts, knowledge and beliefs. Notably, cultural evolution can be studied in an analogous manner to genetic evolution, showing intriguing similarities (as well as differences). In the study concluding this dissertation, we focus on cultural niche construction, a process by which certain evolving cultural traits form a cultural niche that affects the evolution of other cultural traits. We examine the dynamics of cultural niche construction in a metapopulation (a population of populations), where the frequency of one cultural trait (e.g. the level of education) determines the transmission rate of a second trait (e.g. the adoption of fertility reduction preferences) within and between populations. We formulate the Metapopulation Cultural Niche Construction (MPCNC) model by defining the cultural niche induced by the first trait as the construction of a social interaction network on which the second trait may percolate. Analysis of the model reveals dynamics that are markedly different from those observed in a single population, allowing, for example, different (or even opposing) dynamics in each population. In particular, this model can account for the puzzling phenomenon reported in previous studies, that the onset of the demographic transition in different countries occurred at ever lower levels of development.

These studies clearly demonstrate that the interplay between evolution and lifetime adaptation has a marked effect both on the evolutionary trajectory and on the evolving adaptive mechanisms. These interactions form an integral component of the process that gave rise to the complex living organisms found in nature and should be properly incorporated and carefully considered while studying the evolutionary dynamics of adaptive populations.

Contents

List of Figures

Chapter 1

Introduction

In this chapter we will present some of the related work, background and concepts concerning the studies included in this dissertation. In particular, the motivation and challenge in studying the dynamics of adaptive individuals in evolving populations will be presented, as well as a brief overview of various forms of plasticity and learning which relate to our study. A more detailed introduction to each of these subjects is included in the pertaining chapters.

1.1 Adaptive Individuals in Evolving Populations

The modern synthesis describes the evolutionary process as a change in the frequency of alleles determining heritable traits within a population over the course of generations (Futuyma, 2005). Natural selection acts on the genetic variation in the population (created by mutation, genetic recombination and gene flow), leading eventually to the spread and prevalence of those favorable heritable traits that make an organism better adapted to its environment (Fisher, 1930; Wright, 1932; Haldane, 1932). Assuming that each genetic configuration confers a certain survival and reproduction probability (termed *fitness*) upon the individual that carries it, evolution is often conceived as a simple search process in the genotypic space for optima in the fitness landscape (Wright, 1932). Accordingly, traditional mathematical and computational models of evolution regularly assume a predefined, fixed mapping from genotypes to phenotypes and fitness values (Gillespie, 2004; Hartl, 2000).

However, in reality, many organisms also incorporate various learning and plasticity mechanisms, allowing them to better acclimate to the environment dur-

ing their lifetime (Pigliucci, 2001). Developmental plasticity, individual learning and social learning are just a few examples of such mechanisms. The outcome of these learning processes depends not only on the genetic material of the organism, but also on the environment it encounters during its lifetime, the state of other members of the population, and other stochastic parameters involved in learning. Although such acquired traits are not directly inherited in a pure Darwinian framework, clearly, the selection of adaptive individuals is not governed only by their *innate*, genetically encoded, capabilities, but ultimately, by their overall success after learning takes place. Such lifetime learning mechanisms may hence have a major effect on the evolutionary dynamics and may dramatically change the course of evolution.

Accordingly, the hypothesis underlying the studies presented in this dissertation is that the interplay between evolution and lifetime adaptation both markedly affects the dynamics of the evolutionary process and significantly influences the evolving adaptive mechanisms (Parisi and Nolfi, 1996). We believe that without a comprehensive understanding of these interactions, the study of either evolutionary dynamics or learning mechanisms is bound to be lacking. A framework that incorporates the variety of adaptive mechanisms found in nature can shed new light on the ways these mechanisms evolve, operate and interact.

Although the foundations of the research of the interaction between evolution and learning date back to the late 19th century (Baldwin, 1896; Morgan, 1896), the issues we address in this dissertation are presently at the forefront of theoretical and computational research (as demonstrated in the following chapters). The Evolutionary Computation and Adaptive Behavior communities, as well as traditional population biologists, continuously study the dynamics of such interactions in a wide range of environments, tasks and applications (Belew and Mitchell, 1996; Weber and Depew, 2003).

1.2 From Phenotypic Plasticity To Culture

Lifetime adaptation in nature takes many forms, ranging from simple phenotypic (or developmental) plasticity, through individual learning (e.g., via trial and error), to social learning. These learning schemes significantly differ in their underlying mechanisms, level of complexity, capabilities and applications (Heyes and Galef, 1996). In the studies included in this dissertation we touch upon sev-

eral lifetime adaptation mechanisms, and examine the interaction between them and the evolutionary process. Specifically, we focus on three such mechanisms:

1. Individual learning (and, more broadly, phenotypic plasticity)
2. Imitative learning
3. Cultural learning

Although adaptation mechanisms cannot always be clearly categorized or partitioned into a well-defined hierarchy, the above three mechanisms manifest, at least to some extent, an increasing level of complexity in the dynamics they yield when interacting with the evolutionary process: While individual learning is controlled mostly by the interaction between the individual and its environment, imitative learning is largely influenced by the state of other individuals in the evolving population. Similarly, when imitation is frequent and takes place also across generations, cultural evolution becomes relevant and the knowledge possessed by the evolving population as a whole develops and demonstrates essentially independent evolutionary dynamics (Cavalli-Sforza and Feldman, 1981).

1.2.1 Phenotypic Plasticity and Individual Learning

Phenotypic plasticity is defined as the ability of an organism with a given genotype to change its phenotype in response to changes in the environment (Pigliucci, 2001). This ability is often expressed as *norms of reactions*, mapping a range of environments to a range of phenotypes (Schlichting and Pigliucci, 1998). Different organisms display different capacity for phenotypic plasticity, indicating that phenotypic plasticity itself can evolve and has a significant adaptive value (de Jong, 2005).

Traditionally, the notion of phenotypic plasticity is restricted to developmental processes, physiological and behavioral shifts, or environment-dependent gene expression (Dewitt and Scheiner, 2004). However, adopting a broader definition, plasticity can be conceived as any beneficial response to the environment exerted by an individual in order to increase its effective fitness. According to this definition, various learning schemes that allow an individual to better adapt to its environment can be also categorized as phenotypic plasticity mechanisms. On the other hand, according to some researchers, phenotypic plasticity mechanisms

may also take the form of increased random phenotypic variation or developmental noise (rather than a directional beneficial shift) in response to environmental fluctuations (Gavrilets and Hastings, 1994).

The common property of all these mechanisms is the ability of a given genotype to realize certain phenotypes that would not have been accessible without plasticity. Such phenotypic modifications clearly affect the structure of the fitness landscape governing evolution and change the evolutionary dynamics. Hence, in Chapter 2 of this dissertation, we group together the above strict notion of phenotypic plasticity with simple learning mechanisms and random phenotypic variation in our study of their effect on the evolutionary process.

1.2.2 Imitation and the Mirror System

Imitation is an advanced cognitive learning mechanism whereby an individual observes another individual's behavior and replicates it. It is an effective and robust way to learn new traits by utilizing the knowledge already possessed by others. Recent years have seen a renewed interest in imitation (Prinz and Meltzoff, 2002) and it has been addressed by researchers of developmental psychology, social cognition, neurophysiology and neuropsychology (Meltzoff, 1996; Bargh, 1997; Rizzolatti et al., 1996, 2002; Gallese et al., 1996). The definition of imitative behavior in the literature ranges from very strong definitions, which require an aspect of novelty in the imitated behavior, to very weak definitions with no clear boundaries with forms of non-imitative social learning (Billard and Dautenhahn, 1999). Researchers of imitation in humans and animals often categorize social learning into stimulus enhancement, response facilitation, emulation, and true imitation (Byrne and Russon, 1998). Here, we adopt a relatively broad definition and refer to any scenario wherein the behavior of one individual in the population is used to train another individual, as imitation.

A large portion of this dissertation (Chapter 3 and Chapter 4) examines learning by imitation and its interplay with the evolutionary process. There are several major motivations for studying imitative learning: First, it's a highly complex learning scheme, involving numerous cognitive processes. Only recently, work in neuroscience has begun to reveal the mechanisms underlying imitation in the brain (Gallese et al., 1996; Rizzolatti et al., 2002). It is thus intriguing to examine how these complex learning devices emerged. Second, in contrast to individual learning, the success of an imitative individual largely depends on the state of

other members of the population. Hence, in the context of the interaction between learning and evolution, imitative learning gives rise to complex hybrid dynamics which are challenging to study. Furthermore, from an evolutionary computation standpoint, this form of learning is naturally applicable and can be harnessed to improve evolutionary search algorithms where other forms of learning are not applicable (Borenstein and Ruppin, 2003a). Third, learning by imitation is a common and powerful tool in human learning. Researchers of imitation have emphasized the influence of imitation in infants and highlighted imitation as an important channel for learning in school-age children (Meltzoff, 1988). Human beings are by far the most imitative creatures, however, evidence for imitative behavior in other species is continuously accumulating (Kawamura, 1963; Whiten and Ham, 1992; Heyes and Galef, 1996; Whiten et al., 1999). And finally, imitative learning is a prerequisite for cultural evolution (Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981), a process which is further studied in this dissertation.

Focusing on the neuronal devices underlying the capacity to imitate, research in neurophysiology and neuropsychology has led to the exciting discovery of *mirror neurons*. These neurons, found in area F5 in monkeys, discharge both when the monkey performs an action and when it observes another individual making a similar action (Rizzolatti et al., 1996). Studies in humans (using TMS, MEG, and EEG) have revealed an analogous mechanism, whereby motor centers resonate during movement observations (Fadiga et al., 1995; Hari et al., 1998; Cochin et al., 1998; Iacoboni et al., 1999; Buccino et al., 2001). These neuronal devices demonstrate an internal correlation between the representations of perceptual and motor functionalities, and may hence form one of the underlying mechanisms of imitative ability. It is reasonable to assume that imitation has been selected by evolution because individuals capable of imitative learning could have utilized a wider arsenal of favorable acquired behaviors (Whiten and Ham, 1992). Accordingly, our hypothesis is that neuronal structures and mechanisms involved in imitative behavior, such as the mirror system, should be studied and can be explained best by examining the evolutionary process that produced them.

1.2.3 Culture and Cultural Evolution

Imitation, and more generally, social learning, is also the force that drives cultural evolution - the process by which ideas, thoughts, knowledge and beliefs spread

and change over time (Boyd and Richerson, 2005). In the mid-20th century, social sciences started to study how and why various ideas spread in a population. Research concerning the diffusion of innovations was pioneered by Everett Rogers (Rogers, 1962), examining attributes of innovations and adaptation, categories of adopters, diffusion networks and the consequences of innovations. In 1976, Richard Dawkins (Dawkins, 1976) coined the term "meme" to describe a cultural information unit (or cultural gene) that replicates and propagates from one mind to another. The concept of memes as evolutionary replicators was further promoted by Susan Blackmore (Blackmore, 1999) and the founding of memetics - a new scientific discipline that studies evolutionary models of information transmission (Moritz, 1990).

Indeed, cultures never stand still - they evolve, showing many similarities to biological evolutionary processes. Ideas (or for that matter, memes), percolate as units of cultural information, which is analogous in many ways to the spreading of genes, the units of genetic information. Concepts of natural selection and mutation can be also applied to the study of cultural evolution, to explain why one idea becomes extinct while other ideas survive, spread and change over time (Lynch, 1999; Brodie, 1995). These similarities led to the formation of various mathematical frameworks (mostly originating from classical population biology theory) that study cultural transmission and its interaction with the evolutionary process (e.g Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981; Laland, 1992; Feldman and Laland, 1996). These frameworks allow social scientists to model cultural transmission and evolution, using a well-established, quantitative, and rigorous toolbox.

1.3 Studying Evolution and Learning

As evident from the previous sections, the interplay between evolution and learning embodies complex dynamics that consequently render its study a challenging task. A large body of work in recent years has studied the interaction between lifetime learning and genetic evolution. Hinton and Nowlan (1987) introduced a simple model that demonstrates how learning can guide and accelerate evolution. Nolfi et al. (1994) presented experimental results supporting this view, even when the learning task differs from the evolutionary task. Other researchers (Nolfi and Parisi, 1997; Floreano and Mondada, 1996) studied the interaction be-

tween learning and evolution in robots and artificial agents systems. New studies of this basic interaction have been also presented very recently (Paenke et al., 2006; Mills and Watson, 2006). For a collection of classic and recent studies, see Belew and Mitchell (1996). A more multidisciplinary collection can also be found in Weber and Depew (2003). Learning by imitation has also been studied in the context of its interaction with evolution by researchers of artificial intelligence and robotics (Hayes and Demiris, 1994; Billard and Dautenhahn, 1999; Dautenhahn and Nehaniv, 2002b), and in the context of gene-culture co-evolutionary models (Feldman and Laland, 1996). A more detailed overview of related work can be found in each of the pertaining chapters. We believe that the studies presented in this dissertation fill several important gaps in this body of work and form an important contribution to the understanding of these fundamental processes.

1.3.1 A Bidirectional Interaction

As noted before, we believe that the hybrid adaptation dynamics, combining evolution and lifetime learning, have a marked effect both on the evolutionary process and on the evolving learning mechanisms (Parisi and Nolfi, 1996). In particular, in our study of imitative learning we wish to examine two basic questions: (i) Can imitation enhance or accelerate the evolutionary process? (ii) How could the mechanisms underlying imitative learning have evolved and prevailed in the first place? To address these questions and gain a comprehensive understanding of these phenomena we utilize a two-pronged approach:

(a) Study the evolutionary benefits of learning by imitation.

(b) Study the emergence of imitative behavior mechanisms.

The main difference between these two approaches is the underlying assumption: In the first approach, the ability (and incentive) to learn or imitate is assumed to be instinctive, and the focus is on the effect of this mechanism on evolution. We regard learning as a "black box", ignoring the perceptual and neuronal mechanisms it may require. In the second approach, the effect of learning on evolution is assumed to be beneficial, and the emergence of the mechanisms that support imitative behavior is in the center of attention.

1.3.2 Methods and Tools

In this dissertation we present several studies that concern the interplay between evolution and learning. To this end, we apply a plethora of tools, ranging from mathematical analysis to autonomous agent simulations. As is often the case in these studies, only simple and somewhat naive models can be fully analyzed mathematically. Such simple models, however, still offer valuable insights concerning the underlying dynamics and provide important intuition for the construction of more complex models. When a mathematical analysis is not feasible, numerical simulation or an agent-based study is applied.

In our first study, we formulate the process of learning as a transformation that operates on the fitness landscape and modifies the selection pressures. Using this framework we can apply random walk theory (Spitzer, 2001) to provide an accurate measure of evolutionary rates with and without learning and study which forms of learning are beneficial and under which conditions.

In our next two studies, focusing on imitative learning, such a mathematical framework is not possible due to the complexity of the interplay and we hence resort to evolutionary autonomous agent (EAA) simulations (Ruppin, 2002). We first demonstrate how imitative learning can enhance autonomous agents evolution. Aside from the biological implications of this study, it also suggests a simple and powerful method to improve agent-based search algorithms. Second, we present a more complex model based on evolutionary *adaptive* agents (Floreano and Urzelai, 2000) that are capable of imitative learning. Evolving agents that embody simple and fully accessible imitative mechanisms form a simple, yet biologically plausible model for imitation in natural systems. The emergence of unique neuronal devices in artificial neural networks, and a rigorous analysis of these devices' activity and link with the rest of the network provide new insights regarding the evolution and function of imitative learning.

Once the evolutionary origins and effects of imitation have been examined, we turn to study the dynamics of cultural evolution in a metapopulation. Owing to the similarity between genetic and cultural evolution mentioned above, we can again apply a mathematical formulation drawn from traditional population biology theory, and combine it with social network theory for studying the process of cultural niche construction (Ihara and Feldman, 2004; Kendal et al., 2005). The combination of these two theories produces an intriguing model that can account for puzzling phenomena reported in previous social science studies.

Chapter 2

The Effect of Phenotypic Plasticity on the Rate of Evolution in Multipeaked Fitness Landscapes

Based on:

Elhanan Borenstein, Isaac Meilijson and Eytan Ruppin
The effect of phenotypic plasticity on evolution in multipeaked fitness landscapes, *Journal of Evolutionary Biology*, 19(5), 1555-1570, 2006.

Although traits acquired during the lifetime of an organism are not directly inherited in a pure Darwinian framework, they may change the individual's fitness and consequently dramatically alter the dynamics of the evolutionary process (e.g., via genetic assimilation of initially acquired traits through the Baldwin effect, Baldwin, 1896; Morgan, 1896). The interplay between evolution and phenotypic plasticity (e.g., lifetime learning, developmental plasticity, etc.) is thus far from trivial and has been the subject of numerous biological (Waddington, 1942, 1953; Mery and Kawecki, 2002, 2004) and computational (Hinton and Nowlan, 1987; Maynard-Smith, 1987; Belew, 1990; French and Messinger, 1994; Gruau and Whitley, 1993; Menczer and Belew, 1994; Mayley, 1996; Moriarty and Mikkulainen, 1996; Parisi and Nolfi, 1996; Floreano and Urzelai, 1998; Ancel, 1999; Nolfi and Floreano, 1999; Weber and Depew, 2003) studies. Clearly, the capacity of a phenotype to better adjust to its environment has an advantageous

effect in a non-stationary environment, allowing individuals to acclimate to rapid changes that cannot be tracked by the slow evolutionary process (Littman and Ackley, 1991; Todd and Miller, 1991; Nolfi and Parisi, 1997). However, it has been argued that phenotypic plasticity may also be beneficial in static (or slowly changing) environments (Hinton and Nowlan, 1987; Nolfi and Floreano, 1999), facilitating the evolutionary search by smoothing the fitness landscape. Yet, this beneficial effect of phenotypic plasticity on evolution in a static environment is controversial, owing mainly to inconsistent results obtained in various simulation studies and the lack of a rigorous mathematical analysis. We hence focus in this chapter on a mathematical analysis of the interplay between plasticity on evolution, characterizing the plasticity schemes and conditions under which phenotypic plasticity has an advantageous effect and identify its origins.

The remainder of this chapter is organized as follows. We first review previous work studying the effects of phenotypic plasticity on the rate of evolution and identify the source of the discrepancy concerning these effects. In Section 2.2, we use Random Walk (RW) theory to analyze a simple model of evolution and derive a rigorously quantitative measure of evolutionary rate on any given multipeaked landscape. We then turn, in Section 2.3, to examine the effect of plasticity on evolution: We introduce the concept of *innate* vs. *effective* fitness landscapes, representing the effect of plasticity as a transformation, replacing the *innate* fitness landscape that governs selection when no plasticity is present with an alternative *effective* fitness landscape. Using this concept and the derived RW measure, the effect of phenotypic plasticity on evolution can be quantified by comparing the evolutionary convergence rate using the innate vs. the effective fitness functions for selection. We study various plasticity schemes by examining the effective fitness landscapes they induce and their effect on the evolutionary rate. In particular, we examine both deterministic and stochastic models of learning as well as a simple model of random phenotypic variation, and investigate the influence of varying plasticity rate. We further extend the random walk model in Section 2.4 to consider more complicated evolutionary dynamics and better align it with traditional population biology models. The chapter concludes with a discussion of the implications of our findings. The theory and results presented in this chapter have been published in Borenstein et al. (2006c).

2.1 Previous Work

In recent years, a number of researchers have studied the complex interaction between phenotypic plasticity (and specifically, learning) and evolution, employing a variety of methodologies. Comprehensive theories that can combine the two paradigms of evolution and phenotypic plasticity have been recently constructed (Schlichting and Pigliucci, 1998; West-Eberhard, 2003), demonstrating the importance of development and phenotypic response to environmental stimuli in evolutionary theory. In this chapter, however, we focus on one specific question concerning adaptive evolution of phenotypes, examining the effect of plasticity on the convergence rate of the evolutionary process. In the seminal work of Hinton and Nowlan (1987), a simple computational model was introduced to demonstrate how learning can guide and accelerate evolution (see also Maynard-Smith, 1987). Despite its obvious limitations (Nolfi and Floreano, 1999), Hinton and Nowlan's model has successfully demonstrated a distilled model of this effect, bringing the interaction between learning and evolution back to the forefront of scientific research. A large body of work that followed Hinton and Nowlan's study (Belew, 1990; Littman and Ackley, 1991; Gruau and Whitley, 1993; French and Messinger, 1994; Menczer and Belew, 1994; Littman, 1996; Nolfi and Floreano, 1999; Dopazo et al., 2001) further explored the beneficial effect of learning on evolution. Using various simulations of the evolutionary process, these studies demonstrated the benefit of combining learning and evolution in a wide range of stationary and non-stationary environments. Specifically, it has become the accepted wisdom that lifetime learning accelerates evolution in stationary environments by smoothing the fitness landscape and setting up favorable selection preferences for those individuals whose genotypic configurations are in the vicinity of the optimal genotype.

However, even with this ever-growing body of evidence for the advantageous effect of phenotypic plasticity on evolution, rigorous theoretical analysis of this interaction is still scarce. Moreover, such analyses have often found that learning hinders evolution, leading to contradictory predictions (Mery and Kawecki, 2004): Fontanari and Meir (1990) performed a quantitative analysis of an asexual version of the Hinton and Nowlan's model, based on a classic population genetic approach. Corroborating the claims made by Hinton and Nowlan, they showed that learning contributes to the robustness of the evolutionary process against high mutation rates. Studying a more general selection scenario and considering

a one-dimensional Gaussian fitness function, Anderson (1995) found that while learning does have an obvious beneficial effect in changing environments, the advantage of learning in a fixed environment is transient. Representing lifetime acclimation as an increase in the variance of selection and using quantitative genetic models, he showed that learning actually slows the final convergence of the population to a maximal fitness solution. Ancel (2000) further demonstrated that when an extreme fitness scenario is not assumed, phenotypic plasticity does not universally accelerate evolution. Ancel's findings suggest that the Baldwin expediting effect (the term she used for this beneficial effect of learning) may thus not be sufficient to account for the evolutionary success of learning.

The findings of these analytical studies clearly disagree with the beneficial effect of phenotypic plasticity that has been demonstrated in the simulation studies cited above, leading to a long-standing debate. We believe that the source of the discrepancy lies in the structure of the fitness landscapes analyzed. While most of the simulation studies explored relatively complex artificial environments, such that induce highly irregular fitness landscapes, the mathematical anaylses have employed common population dynamics models, focusing on unimodal landscapes (Table 1). These relatively simple landscapes lack one of the key characteristics influencing the convergence rate of the evolutionary process - multiple local optima. The existence of multiple optima (and consequently, multiple domains of attraction) significantly slows down the evolutionary process and hence may make the effect of phenotypic plasticity (or any other mechanism that smooths the landscape) more important. Furthermore, complex genotype-phenotype mapping, developmental processes, epistasis, multiobjective optimization and frequency

Table 2.1: *Findings concerning the effect of phenotypic plasticity on evolution*

Fitness landscape	Simulations results	Analytical results
Extreme[†]	accelerates evolution[1]	accelerates evolution[2]
Unimodal	slows evolution[3]	advantage is transient[4] slows evolution [5]
Multipeaked[‡]	accelerates evolution[6,7,8,9,10] improves evolving solution[8,9,10]	N/A

[†] *One optimal phenotype*, [‡] *Complex environments*, [1] *Hinton and Nowlan 1987*, [2] *Fontanari and Meir 1990*, [3] *Dopazo et al. 2001*, [4] *Anderson 1995*, [5] *Ancel 2000*, [6] *Littman and Ackley 1991*, [7] *Nolfi and Floreano 1999*, [8] *Gruau and Whitley 1993*, [9] *French and Messinger 1994*, [10] *Littman 1996*

dependent selection may all render multiple optima genetic solutions (Wright, 1932), making such multipeaked landscapes a feasible model for biological landscapes and the subject of numerical (Kauffman and Levin, 1987) and experimental (Macken and Perelson, 1989; Korona et al., 1994; Burch and Chao, 1999; Lenski et al., 1999) studies. One noticeable piece of evidence for this characteristic of biological fitness landscapes is demonstrated by a recent study of laboratory evolution of *Escherichia coli* (Fong et al., 2005). Using parallel, replicate adaptive evolution experiments and examining the evolution endpoints, it was shown that the fitness landscape includes distinct peaks of increased adaptive fitness. As shown by Table 1, a rigorous analysis of the effect of phenotypic plasticity on evolution in such landscapes is lacking.

To fill this gap, we focus next on a mathematical analysis of the effects of phenotypic plasticity on evolution in arbitrary multipeaked landscapes. Our analysis facilitates a quantitative comparison between the evolutionary rate with and without phenotypic plasticity, and permits us to identify the origins of the advantageous effect of plasticity in such biologically plausible fitness landscapes.

2.2 Mathematical Analysis of Evolutionary Rate in Multipeaked Landscapes

2.2.1 One-Dimensional Arbitrary Multipeaked Landscapes

Analyzing the dynamics of an evolutionary search on a given landscape has attracted considerable attention in recent years (Kallel et al., 2001). Most efforts focused on studying the geometric properties of fitness landscapes, including multimodality (Goldberg, 1989), autocorrelation (Weinberger, 1990), and neutrality (Huynen et al., 1996a) and on strictly uphill adaptive walk dynamics (Kauffman and Levin, 1987), in an attempt to predict the difficulty of the search task (Stadler, 1995). Here, we provide a direct estimate of the time it will take a stochastic evolutionary process to reach the global optimum on an arbitrary landscape. To obtain a rigorous mathematical analysis of the evolutionary process's dynamics we employ a canonical, one-dimensional model of asexual evolution in a fixed arbitrary environment. Each genotype is encoded by single integer value x, whose fitness value is given by $F(x)$. We assume that the genetic configuration in the first generation is 0 and let N denote the location of the global optimum. Evolution is represented as a simple random walk (RW) process wherein the prob-

abilities p_i (taking a $+1$ step) and $q_i = 1 - p_i$ (taking a -1 step) for each location i are determined according to the differences between the fitness value of i and those of its neighboring genetic configurations (see also Appendix B). It should be noted that in contrast to genetic drift processes that are often modeled by simple *symmetric* random walks (i.e., $p_i = q_i = 1/2$ for every i), here, the $+1$ and -1 step probabilities depend on the fitness landscape structure and hence, p_i and q_i are not necessarily equal and may also vary for different i values. Consequently, we use a *nonsymmetric* random walk model (Spitzer, 2001), allowing us to represent also non-neutral selection schemes. The term *random* thus refers to the stochastic nature of the walk process, where in each point of time the step direction is selected at random with certain probabilities. Such nonsymmetric random walk models are commonly used in physics, engineering, economy and finance (Hughes, 1995).

Within this model, the expected first-passage time from 0 to N, E_0^N, can serve as a good measure for the progress rate of the evolutionary process (describing the time to first encounter of the global optimum) and can be explicitly calculated for any given one-dimensional landscape. Formally, consider a simple RW S_t (± 1 increments) in a changing environment on $\{0, 1, 2, \ldots, N\}$. Let $p_i = P(S_{t+1} = i + 1 | S_t = i)$ and let $q_i = 1 - p_i = P(S_{t+1} = i - 1 | S_t = i)$. Let also ρ_i denote the odds-ratio $\frac{q_i}{p_i}$. Note that $\rho < 1$ indicates a positive selection pressure, $\rho > 1$ indicates a negative selection pressure, and $\rho = 1$ represents regions wherein $p = q$ and thus no selection pressures are exerted (neutral drift). Focusing on the time it takes evolution to reach the global optimum from an initial genetic configuration, it should be noted that the term *positive selection pressure* is used here in the sense that the $+1$ mutant, which is closer to the global optimum than the -1 mutant, is also fitter. Similarly, a *negative selection pressure* indicates regions wherein the -1 mutant is more fit than the $+1$ mutant. Let $p_0 = 1$ and assume that $0 < p_i < 1$ for all $0 < i < N$. As we show in Appendix A.1, the expected first-passage time from 0 to N (i.e., the expected time to first hit N starting at 0) on a given landscape is

$$E_0^N = N + 2 \sum_{\substack{i \leq j \\ 0 < i,j < N}} \prod_{k=i}^{j} \rho_k \ . \tag{2.1}$$

This expression can also be represented as the quadratic form

$$E_0^N = \mathbf{1}'\widetilde{A}\mathbf{1} \tag{2.2}$$

where

$$\widetilde{A} = \begin{pmatrix} 1 & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2\rho_3 & \rho_1\rho_2\rho_3\rho_4 & \cdots & \prod_{k=1}^{N-1}\rho_k \\ \rho_1 & 1 & \rho_2 & \rho_2\rho_3 & \rho_2\rho_3\rho_4 & \cdots & \prod_{k=2}^{N-1}\rho_k \\ \rho_1\rho_2 & \rho_2 & 1 & \rho_3 & \rho_3\rho_4 & \cdots & \prod_{k=3}^{N-1}\rho_k \\ \rho_1\rho_2\rho_3 & \rho_2\rho_3 & \rho_3 & 1 & \rho_4 & \cdots & \prod_{k=4}^{N-1}\rho_k \\ \vdots & & & & & \ddots & \vdots \\ \prod_{k=1}^{N-1}\rho_k & \prod_{k=2}^{N-1}\rho_k & \prod_{k=3}^{N-1}\rho_k & \cdots & \rho_{N-2}\rho_{N-1} & \rho_{N-1} & 1 \end{pmatrix} .$$

As shown in Appendix A.1, this expression for the mean first-passage time from 0 to N on an arbitrary environment, also collapses appropriately to the common expressions in the simple case of a constant environment. Specifically, the above quadratic form yields $E_0^N = N^2$ in a constant symmetric environment (inducing random drift), a linear first-passage time (in N) in a constant nonsymmetric environment where $\rho < 1$ (positive selection) and an exponential first-passage time in a constant nonsymmetric environment where $\rho > 1$ (negative selection).

We further derive an asymptomatic for the first passage-time (see Appendix A.2). For a multipeaked landscape, define the *drawdown* R as the maximal element in the sum above:

$$R = \max_{\substack{i\le j \\ 0\le i,j<N}} \prod_{k=i}^{j} \rho_k \ . \tag{2.3}$$

This characterizing feature of the landscape has also been termed in the literature *gap* or *extent* (Noskowicz and Goldhirsch, 1990; Meilijson, 2003). We show that the expected first-passage time is dominated by R, whereby E_0^N is sharply bounded from above by (see Appendix A for a full analysis):

$$E_0^N \le N^2(\frac{1+R}{2}) \ . \tag{2.4}$$

Furthermore, as demonstrated in Appendix A.2, this bound is reached when all ρ_i values contributing to the drawdown collapse to a single point in the landscape. Hence, crossing the fitness landscape from the initial genetic configuration toward the global optimum, R corresponds to the height difference between the record high fitness value to the consecutive record low fitness value (i.e., the largest descent along this fitness landscape), as illustrated in Figure 2.1.

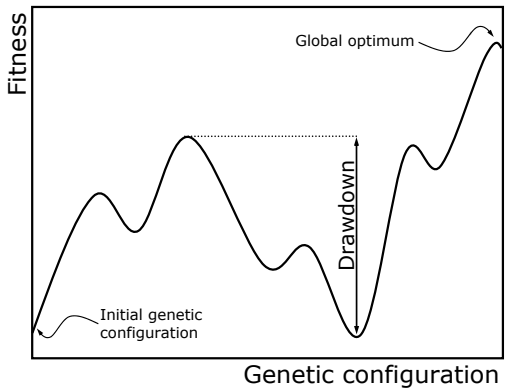


Figure 2.1: *A qualitative illustration of the fitness drawdown.*

2.2.2 Multidimensional Arbitrary Multipeaked Landscapes

While the one-dimensional case allows for a rigorous analysis of expected first-passage times, such an analysis is hard to obtain for a multidimensional fitness landscape. We show, however, that an analogous dominance of the drawdown on the evolutionary rate can be inferred through numerical simulations. In this case, we consider all the possible pathways from the initial configuration to the global optimum configuration. Each of these pathways can be conceived as a simple one-dimensional landscape, with a specific drawdown value. We will term the pathway with the minimal drawdown value 'the *Principal-Pathway*' and the drawdown value it induces 'the *Principal-Pathway drawdown*'. We maintain that the Principal-Pathway drawdown dominates the random walk first-passage time on a given multidimensional landscape in a similar manner to that shown in the one-dimensional case (see Appendix A.3 for details and discussion).

We validate the strong correlation between the Principal-Pathway drawdown and the first-passage time through numerical simulations, using common multimodal benchmark functions. These functions are utilized to generate numerous landscapes with 1, 2 and 3 dimensions and varying drawdown values. The Principal-Pathway drawdown (or simple drawdown value in the one-dimensional case) and the expected first-passage time to the global-optimum, E_0^N, were evaluated for each landscape (see Appendix A.3 for details). As demonstrated in Figure 2.2, there is a strong correlation between the drawdown value in each fitness landscape and the first-passage time of the random walk process, corroborating the validity of the fitness landscape's drawdown as a measurable bound for the expected evolutionary progress rate.

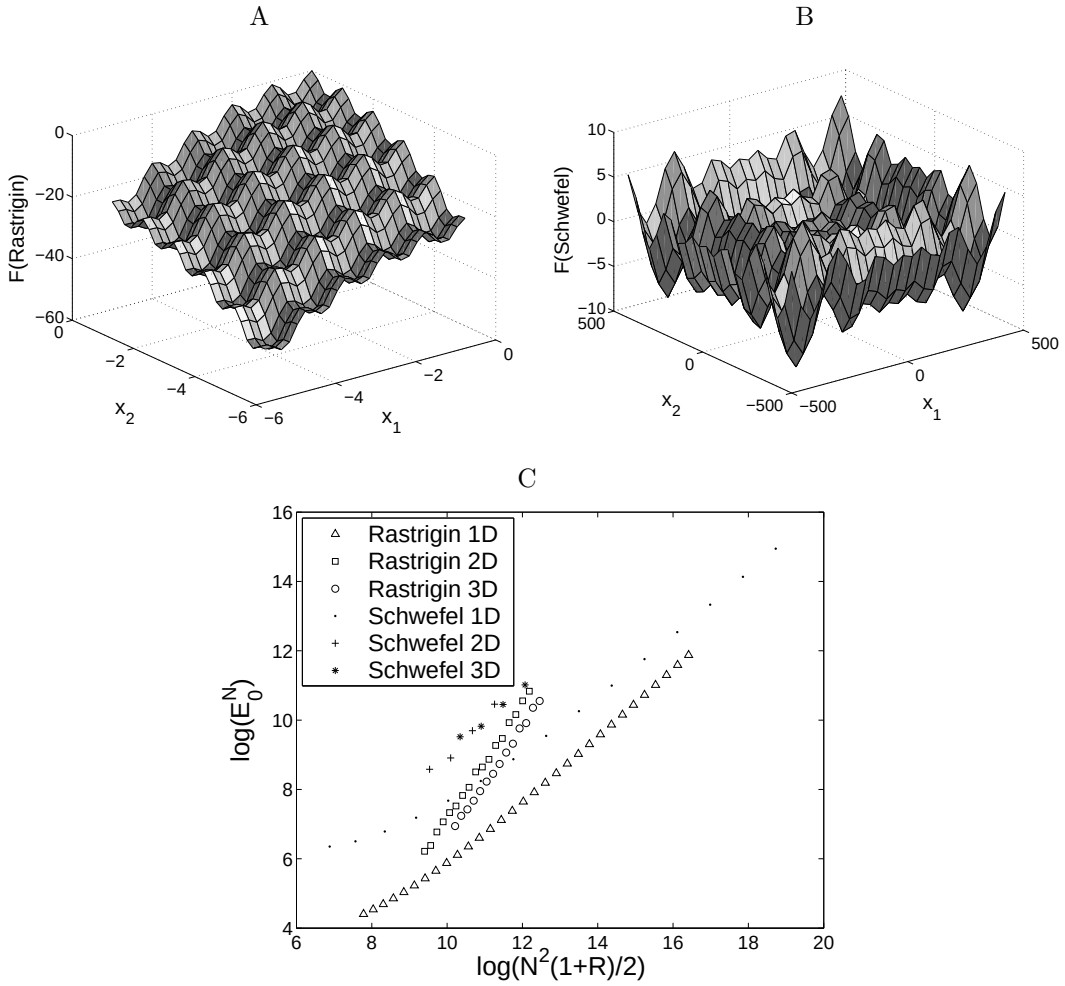


Figure 2.2: *The correlation between the expected first-passage time and the landscape drawdown in multidimensional landscapes.* An example of two-dimensional landscapes generated by the (**A**) Rastrigin function ($C_r = 3$) and the (**B**) Schwefel function ($C_s = 0.01$) (see Appendix A.3 for more details on these functions). (**C**) The correlation between the expected first-passage time and the landscape drawdown. Numerous landscapes of different dimensions and varying drawdown were tested. E_0^N was evaluated for each landscape through 100 random walk simulations. Landscapes for which not all simulations hit the global optimum within 250,000 steps were excluded from our analysis. For the one-dimensional landscapes, the expected first-passage time was analytically calculated as described before. The drawdown value is normalized according to the asymptotic bound found for the one-dimensional case, where N denotes the length of the pathway. As the drawdown values (and consequently, the first-passage times) exhibit a large variation, a logarithmic scale was applied. Linear regression analysis for the correlation between R and E_0^N in the Rastrigin landscapes yielded coefficient of determination values (percent of variance explained) of 0.995, 0.9971 and 0.9961 for the 1, 2 and 3 dimensional landscapes respectively. In the Schwefel landscapes the resulting coefficient of determination values were 0.9809, 0.9718 and 0.9818 for the same dimensions.

2.3 The Effect of Phenotypic Plasticity on Evolution

2.3.1 Innate vs. Effective Fitness Landscapes

Having established a measure of evolutionary rate, we turn to examine the dynamics of the evolutionary process in two modes: In the first, *non-plastic* mode, phenotypic plasticity is absent and the fitness value $F(\vec{x})$ assigned to each genotypic configuration $\vec{x}$ is uniquely determined according to the innate survival and reproduction probability of the phenotype that it encodes, termed *innate fitness*. In the second, *plastic* mode, phenotypes can vary during their lifetime and as a result may effectively gain a different (and potentially higher) fitness value. Clearly, in this mode, selection operates according to the effective fitness value obtained by each individual. We denote this *effective fitness* value, $F_{efc}(\vec{x})$. It should be noted that both the innate and the effective fitness landscapes ultimately correspond to the common notion of fitness in natural selection. We use the terms *innate* and *effective* fitness to simply distinguish between the fitness values that govern selection without and with plasticity. Phenotypic plasticity hence manifests itself as a transformation of the fitness landscape, replacing the innate fitness that initially governed selection with an effective fitness landscape (Figure 2.3A). The strength of this simple model lies in the fact that the complex dynamics of a hybrid process combining evolution and phenotypic plasticity can be studied by examining the simpler dynamics of a pure evolutionary process on the appropriate effective fitness landscape. The effect of phenotypic plasticity on the evolutionary convergence rate can be measured by comparing the time it takes the evolutionary process to obtain an optimal genotype using the innate vs. the effective fitness functions for selection. In particular, the random walk analysis and drawdown value presented in Section 2.2 provide a measurable bound for the expected evolutionary progress rate for *any* given landscape, allowing for a direct quantitative comparison between the innate and the effective fitness landscapes.

In this section we examine several models of phenotypic plasticity, for which the resulting effective fitness landscapes can be explicitly constructed and studied. We start with a simple model of deterministic learning, where phenotypic plasticity is manifested as a hill climbing process (either ideal or partial). Next, we examine a stochastic learning model in which the direction of each learning iteration is determined probabilistically. Finally, although traditionally not de-

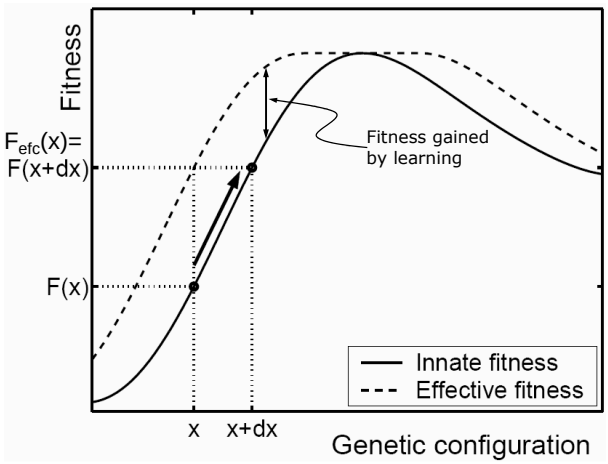
fined as a form of phenotypic plasticity, a model of random phenotypic variation is presented, where each genotype randomly realizes one of several alternative phenotypes. We compare the drawdown of the original, innate landscape with that of the resulting effective fitness landscapes induced by these plasticity schemes. Identifying the schemes and conditions under which the fitness landscape drawdown is reduced (and thus, the rate of evolution is accelerated) provides a characterization of the beneficial effect of phenotypic plasticity.

2.3.2 Ideal and Partial Deterministic Learning

Individual learning, as a form of phenotypic plasticity, is modelled as an iterative process of phenotypic modifications aimed at increasing the individual's effective fitness, taking the form of a simple gradient-ascent process in the genotype/phenotype space. As in previous studies (Hinton and Nowlan, 1987; Anderson, 1995), we focus on the simple case of one-to-one mapping from genotype to phenotype and assume that learning and evolution both operate on the same fixed fitness landscape (see also Section 2.5). We examine a simple model of learning where during each learning episode (iteration), an individual compares the innate fitness value of its current configuration with those of slightly modified configurations, and adopts a modified configuration if the latter's innate fitness value is higher (see Figure 2.3A for an illustration). Such learning iterations may repeat, allowing the individual to adopt behaviors further away from its innate one, resulting in a modification of its effective fitness accordingly (although its genotype remains unchanged).

We first consider an *ideal deterministic learning* model, where each individual repeatedly employs such deterministic hill-climbing learning iterations until it reaches the nearest local-optimum and no further improvement of its effective fitness is possible. As demonstrated in Figure 2.3B, in ideal learning all genetic configurations in the region forming the basin of attraction of a given local optimum will eventually acquire the same effective fitness value, equal to the innate fitness of the local optimum, totally suppressing selection pressures within each such region. In the one-dimensional case, ideal learning transforms each given consecutive pair of descending and ascending intervals in the innate fitness landscape into a single step function in the effective fitness landscape, whose height is equal to the difference between the extents of this descent and the consequent ascent (Figure 2.3B). The drawdown characterizing the effective fitness landscape

A



B

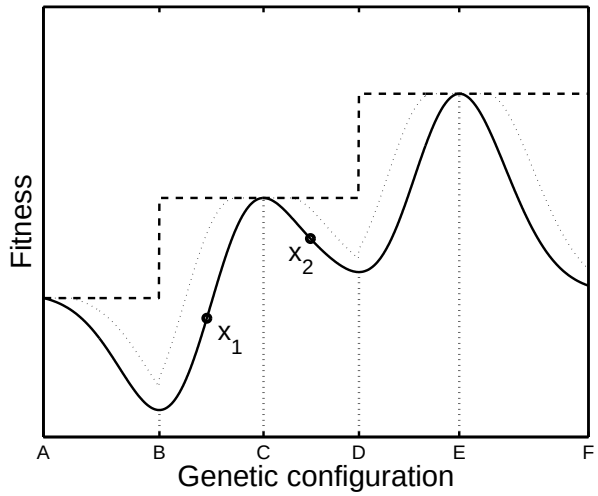


Figure 2.3: *The effect of deterministic learning on the fitness landscape.* (**A**) An individual with genotype configuration x and innate fitness value $F(x)$ may acclimate by learning (illustrated here as a simple gradient ascent process) and gain a fitness value of $F(x + \Delta x)$. As the genotype of this individual remains unchanged, the effective fitness value $F_{efc}(x) = F(x + \Delta x)$ is applied to x. (**B**) The innate fitness function (solid line) and the effective functions obtained with partial learning, i.e., after a limited number of hill-climbing iterations (dotted line), and with ideal learning (dashed line). In the ideal learning scheme all configurations in the basin of attraction of a given local optimum (e.g. genotypes x_1 and x_2 in the interval $[B, D]$) acquire the same effective fitness value, that of the local optimum (C).

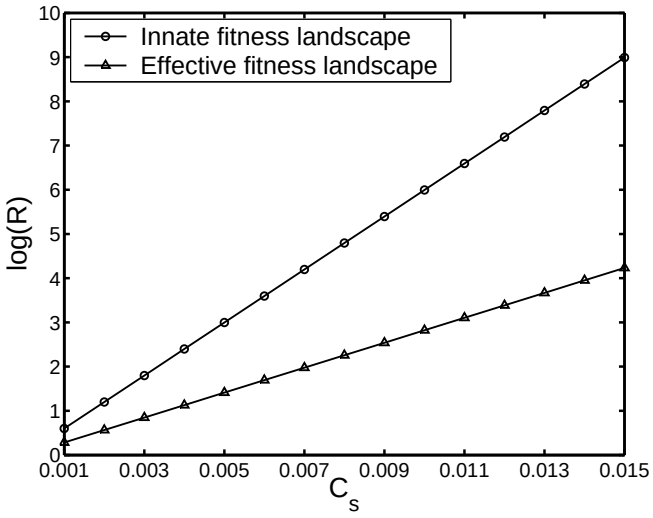


Figure 2.4: *The effect of ideal deterministic learning on the fitness landscape drawdown.* The generalized Schwefel function described in Appendix A.3 is used to generate two-dimensional fitness landscapes with varying ruggedness (tuned via the C_s parameter whose higher values denote increased levels of ruggedness). Ideal learning is then applied to produce the corresponding effective fitness landscape. The Principal-Pathway drawdown calculated for each landscape is illustrated. Evidently, the drawdown induced by the effective fitness is significantly smaller (note the logarithmic scale) than that induced by the innate fitness.

is hence smaller (or equal, in the worst case) than that induced by the original, innate fitness landscape (see also Figure 2.4). As demonstrated above, a smaller drawdown value implies a shorter first-passage time, making the beneficial effect of this learning scheme evident.

The mathematical analysis presented above in Section 2.2 can account for the seemingly contradictory findings of previous studies concerning simple fitness landscapes (Anderson, 1995; Ancel, 2000; Dopazo et al., 2001), in which learning was found to hinder the evolutionary process. While the evolutionary rate in multipeaked landscapes is dominated by the landscape drawdown, in the simple unimodal scenario, no drawdown exists ($R = 1$), and consequently the marked beneficial effect of learning demonstrated above is absent. In terms of our model, in such single peaked landscapes, the evolutionary process is scaled down to a simple random walk within a positive slope interval. Since learning decreases the slope of the fitness function (causing $\frac{q}{p}$ to approach unity), our model clearly shows (see Appendix A.1, Example 2, concerning a constant environment) that learning would slow down the convergence rate, as Anderson (1995) and Ancel

(2000) have indeed found. In particular, learning schemes that cancel all selection pressures and produce a totally flat effective landscape (as is the case in ideal learning) result in a random drift process with quadratic first-passage times, markedly slower than the linear first-passage time in the innate single-peaked landscape (see Appendix A.1, Example 1 and Example 2). It is only in a multipeaked landscape (Wright, 1932; Kauffman and Levin, 1987; Kauffman, 1993; Korona et al., 1994; Burch and Chao, 1999; Lenski et al., 1999; Fong et al., 2005), where the overall evolutionary rate is dominated by the exponential passage time in the negative selection regions, that the beneficial effect of ideal learning is demonstrated.

When learning "resources" are limited (e.g. learning is bounded by a certain cost) and individuals employ only a limited number of hill-climbing iterations, a *partial plastic mode* is obtained rather than ideal learning model. In this mode, not all genetic configurations in the basin of attraction of each local optimum will inevitably gain the same effective fitness value. Individuals with innate genetic configurations farther from the local optimum configuration do improve through learning (and gain a higher effective fitness value), but may fall short of reaching the local optimum's exact fitness level. The effective fitness landscape forms an intermediate state between the plastic and non-plastic modes, including both intervals with constant fitness and intervals with positive or negative slopes (Figure 2.3B). Clearly, a partial learning scheme still reduces the extent of the innate landscape drawdown (and hence, will still accelerate evolution), though it does not cancel them altogether. It is thus expected that this learning mode will yield an intermediate convergence time, progressing slower than the ideal plastic mode, but still faster than the non-plastic one.

These effects of deterministic learning are validated numerically (see Appendix A.3 for the simulation details). The mean first-passage time of each genetic configuration x (i.e., the expected time to first hit x starting at 0) is illustrated in Figure 2.5B. The curves clearly agree with the results of our analysis. Figure 2.5C demonstrates the average **innate** fitness value of the evolving individual as a function of generation of the evolutionary process. Evidently, individuals evolving in the plastic mode converge much faster to the global optimum and gain higher fitness values. Although learning individuals using ideal learning do not converge to the exact global optimum, they successfully reach its basin of attraction and possess higher innate fitness values than non-plastic

individuals who tend to get trapped in remote local optima. Examining the effect of partial learning, we find, as expected, that this mode yields an intermediate convergence time, progressing slower than the plastic mode, but still faster than the non-plastic one (Figure 2.5B). Furthermore, as this form of learning does not entirely suppress the selection pressures in each optimum domain, it allows individuals that hit the global optimum basin of attraction to converge closer to the exact global optimum configuration, resulting in overall better average innate fitness values than those obtained with ideal learning (Figure 2.5C).

2.3.3 Stochastic Learning

Both the ideal and the partial learning schemes examined above embody two basic characteristics: locality and accuracy; learning was assumed to exploit only local information about the fitness landscape on which it operates, and to do so with complete accuracy. It is these two features that guarantee the preservation of extrema domains in the effective landscape. However, the lack of complete environmental data, sensory input noise, imperfect information processing and nondeterministic decision making, all make a stochastic learning process more plausible as a model of learning in biology. Yet, since stochastic local search schemes are not bound to take the steepest ascent route, and can potentially discover remote local optima, the effective fitness function they yield may have a different regional structure than that of the original innate fitness. For example, in an extreme scenario, a stochastic learning algorithm may allow any innate genetic configuration to successfully reach the global optimum solution, totally suppressing genetic selection pressures. In this scenario the evolutionary process turns into a random drift, which, as was demonstrated in Appendix A, yields a quadratic first-passage time.

To study the effects of stochastic learning, we use a simple variation of our model, where the hill-climbing learning algorithm is replaced with a simulated annealing (SA) optimization process (Kirkpatrick et al., 1983) (see Appendix B for details). Clearly, stochastic learning does not guarantee a consistent fitness gain each time learning is applied. Consequently, the effective fitness value assigned to each genetic configuration varies from one learning process to the other and a deterministic effective fitness function cannot be explicitly constructed in advance. Yet, examining the characteristics of the *average* effective fitness function constructed by this stochastic scheme (see, for example, Figure 2.7B), it is

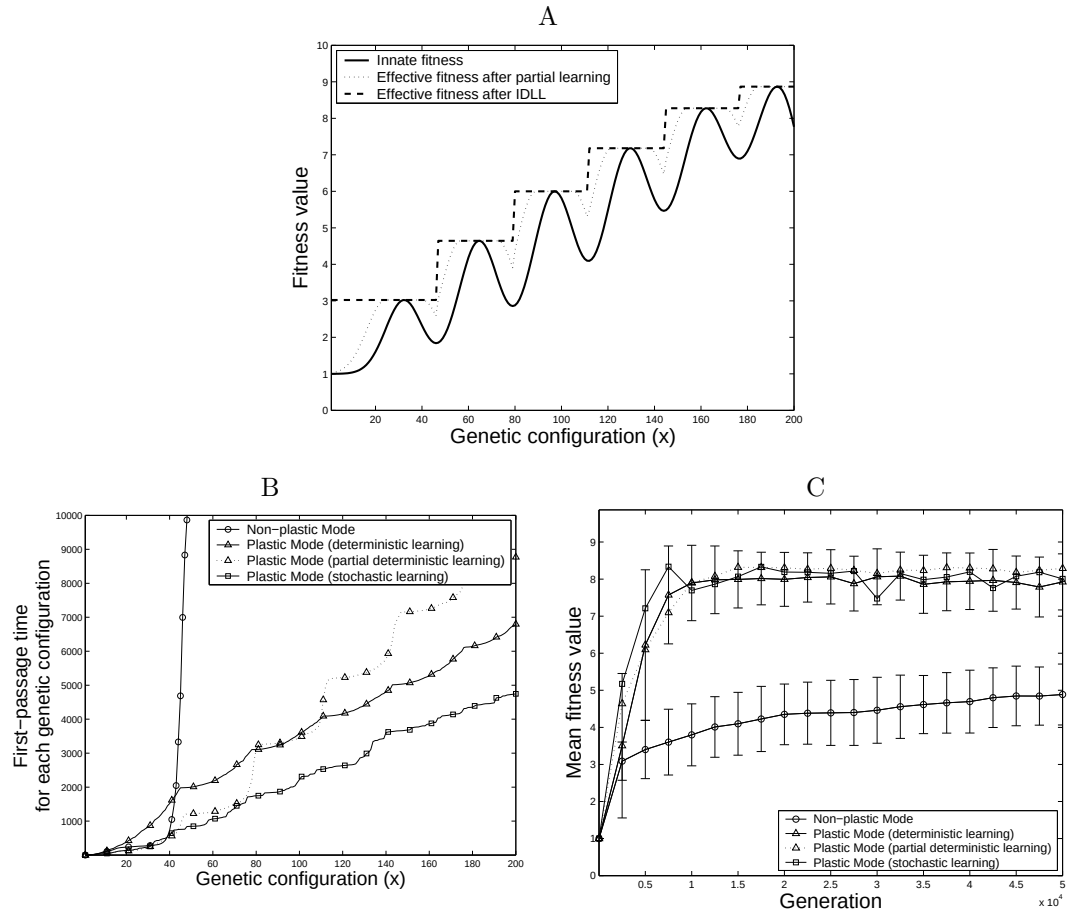


Figure 2.5: *The effect of ideal deterministic learning (IDLL), partial deterministic learning and stochastic learning on the evolutionary process in the one-dimensional case.* (**A**) A one-dimensional innate fitness function was defined on the interval $[1, 200]$ as a sum of several Gaussian functions, yielding a continuous, multipeaked function $F(x)$ (solid line). Various plasticity schemes were then applied to produce the corresponding effective fitness functions (see Appendix B for more details). (**B**) The first-passage time of each genetic configuration x (i.e., the average time to first hit x). Each curve represents the average result of 100 runs for the deterministic learning simulation and 10 runs for stochastic learning simulation. In the plastic mode, using a deterministic learning scheme, all 100 simulation runs hit the global optimum within less than 16,200 generations. In the non-plastic mode, although the linear expected first hitting time in the positive slope intervals yields a fast progress, the exponential behavior in the negative slope intervals dominates the dynamics of the random walk and hinders the evolutionary process. Out of 100 simulated evolutions, each running for a maximum of 200,000 generations, 16 never hit the third local optimum ($x = 97$) and 67 failed to hit the forth ($x = 130$). (**C**) The mean **innate** fitness value as a function of generation. The standard deviation of the non-plastic mode and plastic mode with ideal deterministic learning is also illustrated.

clear that the drawdown induced by the average effective landscape is smaller than that of the original innate landscape.

Numerical simulations of an evolving population applying stochastic learning (where the effective fitness is appropriately evaluated repeatedly for each individual in each generation) validate that this learning scheme indeed accelerates the evolutionary process. Evidently (Figure 2.5B-C), also with a stochastic learning paradigm, learning individuals converge faster and gain significantly higher innate fitness values than those evolving in the non-plastic mode, obtaining values similar to those obtained with deterministic learning. Furthermore, it is shown that stochastic learning not only accelerates evolution in comparison to the non-plastic mode, but yields superior evolutionary convergence rates even in comparison to those obtained in the deterministic learning scheme examined above. The superiority of this scheme can be attributed to the resulting effective fitness landscape which is smoother (on average) than the one produced by a deterministic scheme. This can allow individuals near the boundary between basins of attraction to stochastically converge to either of the two adjacent local optima.

It should be noted that in the extreme case, stochastic learning can produce a totally flat effective fitness landscape, suppressing all selection pressures, even in a multipeaked landscape. However, while the quadratic first-passage time induced by random drift on a flat landscape hinders evolution in comparison to the linear time on a single peaked landscape, it is still superior to the exponential first-passage time expected on a multipeaked landscape.

2.3.4 Random Phenotypic Variation

The learning schemes discussed above represent a directed model of plasticity, aiming at increasing the individual effective fitness. However, phenotypic plasticity may also take the form of increased phenotypic variation (or developmental noise) in response to environmental fluctuations (Gavrilets and Hastings, 1994). Although biologists often refer to phenotypic plasticity as a *beneficial* response to the environment (rather than a random variation), here, as in Ancel and Fontana (2000), we wish to examine the effect of a random phenotypic flexibility scheme . This form of plasticity may be more common in molecular evolution, wherein a certain genotype may realize a range of phenotypic configurations according to the microenvironemntal context.

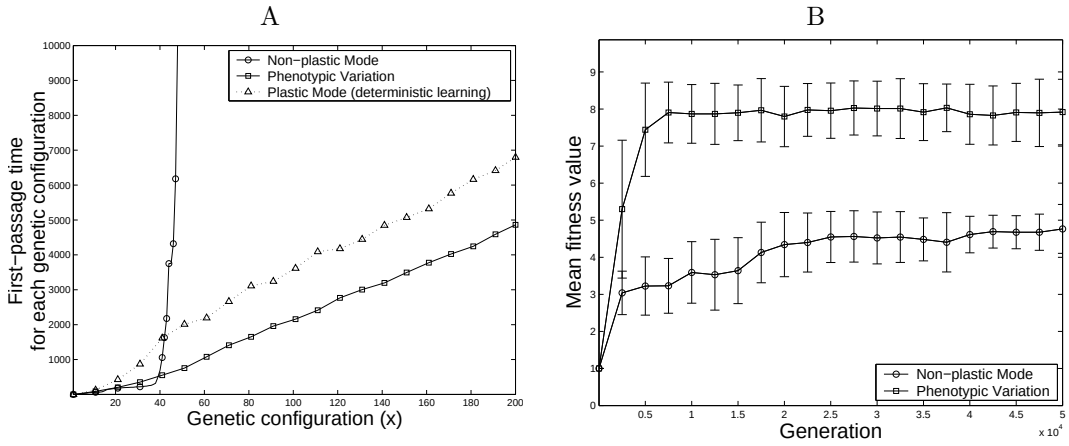


Figure 2.6: *The effect of random phenotypic variation with* $\Delta d = 15$ *on the evolutionary process.* Each curve represents the average result of 100 simulation runs. (**A**) Mean first-passage time as a function of the genetic configuration x. The results obtained under the ideal deterministic learning scheme are depicted for comparison. (**B**) The mean and standard deviation of the innate fitness value as a function of generation.

Such random perturbations are clearly not necessarily in the direction of improved fitness. However, assuming some correlation between the phenotypic and genotypic spaces, on average, genotypes assigned with low innate fitness values will gain more by these perturbations than genotypes with high innate fitness values. In the extreme case, genotypes located in a local minima of the fitness landscape can only gain higher effective fitness by realizing phenotypes of neighboring configurations, while genotypes located on local maxima will inevitably gain lower effective fitness. These dynamics, although stochastic, lead to a reduction in the fitness landscape drawdown, and hence, according to our analysis, accelerate evolution. Applying a simple model of genetic variation, where each genotype "develops" into a phenotype associated with a randomly selected neighboring genetic configuration within a predefined range, Δd (see Appendix B for more details), we validate the beneficial effect of this plasticity scheme (Figure 2.6).

A simple example of such a plasticity scheme can be demonstrate in the RNA secondary structure. While the minimum free energy (MFE) secondary structure of an RNA sequence defines a simple mapping from genotypes to phenotypes, in practice, an RNA molecule may fold into a wide range of secondary structure configurations, providing that the energy barriers are sufficiently small. These

phenomena can be conceived as the RNA equivalent of phenotypic plasticity or developmental noise (Ancel and Fontana, 2000). Furthermore, as demonstrated by Ancel and Fontana (2000), there is a significant correlation between the repertoire of thermodynamically accessible configurations and genetically accessible configurations (a phenomenon they term *plastogenetic congruence*). This correlation implies that RNA molecules that make such thermodynamic transitions can be described as effectively realizing MFE structures associated with neighboring genotypic configurations. Interestingly, examining whether phenotypic plasticity expedites the evolutionary discovery of new structures in RNA, Ancel and Fontana (2000) find that no such expediting occurs due to intrinsic properties of the RNA genotype-phenotype map. Specifically, the high neutrality incorporated in the RNA genotype-phenotype map and its organization make the beneficial effect of plasticity restricted to relatively small regions in the genotypic space. Consequentially, the benefit gained by plasticity is negligible compared to the time it takes to discover these regions in the first place.

2.3.5 Varying Learning Rates

Evidently, different plasticity schemes yield different dynamics of the evolutionary process and result in different convergence rates. The number of phenotype acclimation iterations employed during life or the phenotypic variation range (which will both be referred to here as the *plasticity rate*) may also influence the convergence rate and the stability of the evolutionary process, as was demonstrated by the favorable effects of the partial learning scheme. To further explore and compare the effect of these learning schemes and in particular the effect of varying plasticity rates, an additional set of simulations was carried out. Figures 2.7A-C illustrate the effective fitness functions constructed by each scenario that was tested. To reduce the long computation time in stochastic learning simulations, the *mean* effective fitness was used as a constant effective fitness landscape, approximating a genuine stochastic learning paradigm.

Two measures were examined for each scenario: The *overall convergence time*, which was taken as the first-passage time of the global optimum ($x = 193$), and the *genetic stability* of the evolving individuals that was measured as the average genetic deviation from the global optimum configuration throughout 1000 generations following the first-passage time. As shown in Figure 2.7D, the best convergence time for deterministic learning is obtained with 10 learning iterations.

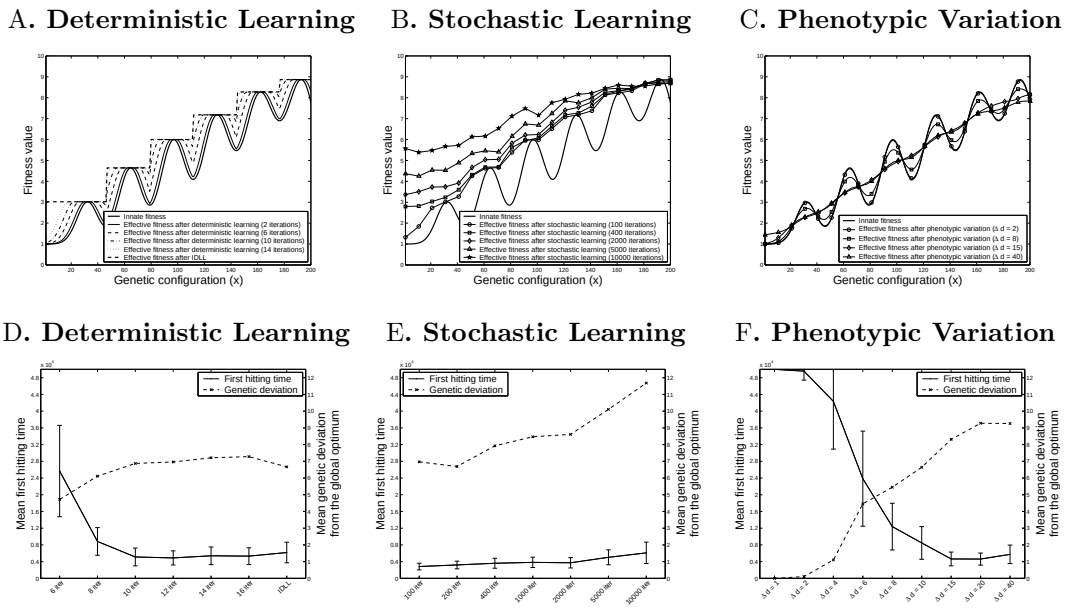


Figure 2.7: *The effect of various plasticity schemes and varying plasticity rate on evolution.* (**A-C**) The average effective fitness functions resulting by varying numbers of deterministic and stochastic learning iterations and by varying phenotypic variation range. Employing more than 14 deterministic learning iterations results with an effective fitness similar to the one obtained by ideal learning. The curves illustrated in Figure C represent the average of 50,000 runs. (**D-F**) The average convergence rate, measured as the mean first-passage time of the global optimum (solid line) and genetic stability (dashed line) obtained for varying plasticity rates. Most simulation runs using less than 6 deterministic learning iterations did not converge to the global optimum.

There is also a clear tradeoff between the convergence time and the genetic stability of the resulting evolutionary process. Figure 2.7E illustrates the results for stochastic learning schemes with a varying number of SA iterations. Evidently, a low number of stochastic learning iterations results in faster convergence rates than those obtained with deterministic schemes and still yields relatively stable genetic solutions. Only when the learning process employs a considerably large number of stochastic iterations it diverges from the original structure of the innate fitness function (see Figure 2.7B), reducing dramatically the evolutionary selection pressures and consequently reducing the genetic stability of the evolving individuals. A clear tradeoff between the convergence time and the genetic stability is also demonstrated in the phenotypic variation experiments (Figure 2.7F). Applying a large variation (e.g., $\Delta d = 15$) results in a fast convergence rate, comparable to that obtained by deterministic learning, but dramatically reduces the genetic stability.

2.4 Numerical Extensions of the Random Walk Model

While the random walk model presented above allows for a rigorous analysis of first passage times, there are a few extensions that make this model more biologically plausible. In particular, we wish to examine whether the possibility to stay in the same genetic configuration over several generations and the transformation from the fitness landscape values to fixation probabilities affect the resulting dynamics. The analytical treatment of the behavior of the RW model with these extensions turns to be a difficult challenge, and hence these extensions are examined numerically in this section.

2.4.1 Random Walk with Static Periods

As the RW model presented above assumes that in each point in time the walk process takes either a $+1$ or a -1 step, our measure of first passage time can be interpreted as measuring the number of mutation events required to reach the global optima. However, in practice, the population may stay in the same genetic configuration for many generations. Hence, if we wish to measure the convergence time in number of generations, we should also allow the walk process to stay in the same configuration with some probability.

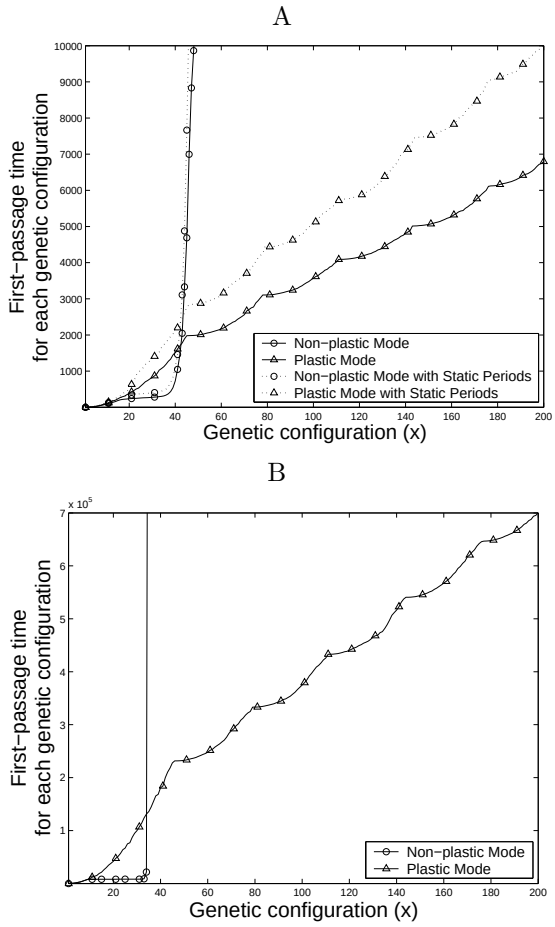


Figure 2.8: *Numerical extensions of the random-walk model.* (**A**) First-passage times for each genetic configuration using a random walk model with static periods. (**B**) First-passage times for each genetic configuration using Kimura's fixation probabilites.

As demonstrated in Figure 2.8A, introducing this extension to the model (see Appendix B for the simulation details) does not significantly change the resulting dynamics, and in particular, preserves the superiority of the plastic mode vs. the non-plastic mode. Apparently, although the probability to stay in the same genetic configuration is higher in neutral regions of the landscapes (which are abundant in the landscape induced by ideal plasticity) than in positive selection regions, the overall convergence time is still dominated by the slow exponential time it takes to cross negative selection regions (and, particularity, the fitness drawdown). The main effect of this extension is manifested by an overall, relatively constant delay in the time to reach each genetic configuration in comparison to the basic model.

2.4.2 Random Walk with Kimura's Fixation Probabilities

To further relax some of the differences between our model and traditional population biology models, we apply Kimura's theory for population dynamics on neutral (or nearly neutral) landscapes (Kimura, 1983). Specifically, we use Kimura fixation probabilities rather than the Boltzmann scaling (see also Appendix B) to determine the RW probabilities. Under this extension, our RW model is closely related to common evolutionary dynamics models on nearly neutral landscapes, taking also the population size into account. Moreover, assuming that the mutation rate is slow (and hence, each mutation becomes either fixated or extinct before the next mutation arises), the first passage time measure, E_0^N, indicates the number of mutations that are required to *appear* until the population fixates on the landscape's global optimum and can be now easily translated to the expected number of generations simply by multiplying E_0^N by the mutation rate ν. As in the previous section, this extension allows for the population to stay in the same genetic configuration over several generations if none of the neighboring mutations is fixate.

As can be seen in Figure 2.8B, using Kimura's probabilities dramatically slows down the entire process (note the first-passage time scale), mainly due to the significantly smaller probabilities of a neutral ($p = 1/2N_e$) or deleterious mutation to fixate. Clearly, in this model, fixation of a slightly beneficial mutation is markedly faster than that of a neutral mutation. Yet, as in our previous analysis, the probability of fixation of a slightly deleterious mutation becomes infinitesimally (exponentially) small as the population size increases and hence still dominates the overall convergence time, resulting in the superiority of the plastic mode over the non-plastic one.

2.5 Discussion

This chapter focuses on the effects of phenotypic plasticity on the evolutionary convergence rate in stationary environments. We use random walk theory to derive a measure for the rate of evolution on arbitrary multipeaked fitness landscapes, and demonstrate that the convergence rate is dominated by the landscape drawdown. Examining various phenotypic plasticity schemes we find that these schemes decrease the landscape drawdown and hence, accelerate evolution. These findings introduce a rigorous quantitative confirmation for the common hypothe-

sis stating that phenotypic plasticity expedites evolution by smoothing the fitness landscape and identify the origins of this phenomenon. Our analysis provides a measure for both the convergence rate bottleneck induced by the landscape drawdown, and the benefit gained by smoothing the landscape and reducing the extent of this drawdown.

Our findings suggest two fundamental principles that affect the interaction between phenotypic plasticity and evolution. First, as the benefit of plasticity stems from its capacity to smooth the fitness landscape, this effect will be revealed only in multipeaked landscapes wherein the evolutionary rate is dominated by the landscape drawdown. Conversely, if selection takes place only within the domain of a simple fitness function that does not include multiple local optima, plasticity hinders the evolutionary process as was also demonstrated in previous studies (Anderson, 1995; Ancel, 2000). Second, our analysis suggests that plasticity has a beneficial effect on evolution when genotypes with low innate fitness values (e.g., individuals at local fitness minima) gain more through phenotypic plasticity than genotypes with high innate fitness values (e.g., with local maxima configurations). These dynamics are governed mainly by the correlation between the genotypic and phenotypic spaces (determined by the genotype-phenotype mapping). As demonstrated in Section 2.3.4, when such a correlation exists, phenotypic modifications due to plasticity are analogous to perturbations in the genetic space, and consequently plasticity yields, on average, a higher gain for those individuals whose genotypes are located on local fitness minima.

Mayley (1997) discussed the effects of learning on evolution in a rugged landscape, arguing that such settings may give rise to two competing effects: A *guiding* effect, helping evolution to detect individuals located near superior local optima, and a *hiding* effect where learning suppresses the selection pressures within each local optima basin of attraction. Our analysis demonstrates that in this scenario, the guiding effect outweighs the hiding effect, resulting with an overall acceleration of the evolutionary process.

Plasticity may also have associated costs, either direct, due to the energy and time it takes to learn a new behavior, or indirect, due to the delay until an individual can exhibit a successful acquired behavior. Learning costs have been largely discussed in the literature of the interaction between evolution and plasticity. Mayley (1996) showed that the selection pressures for acquired traits to become genetically assimilated depend on the trade-off between learning costs

and learning benefits. Examining the evolution of language, Munroe and Cangelosi (2002) demonstrated that when language learning costs are high, agents genetically assimilate explicit properties of the specific language they are exposed to. In the context of our model, associated plasticity costs can clearly affect the resulting fitness gain and consequently, the effective fitness landscape. Assuming a simple model of learning costs, where cost is proportional to the fitness gained by learning (or to the extent of the difference between the innate and acquired configurations), two scenarios should be considered: If the cost of learning is higher than the learning gain, individuals in the plastic mode will in fact have a lower fitness than individuals in the non-plastic mode. Furthermore, individuals at local fitness minima will lose more by learning that those located at local maxima configurations. Hence, in this scenario, the drawdown associated with the effective fitness landscape will be higher than that of the innate landscape, hindering the evolutionary process. However, this form of learning is clearly not adaptive and in most cases would not evolve. In the second scenario, when learning costs are lower than the fitness gained by learning, learning will still decrease the extent of the landscape drawdown (as demonstrated, for example, in the partial deterministic learning model), but this decrease will be smaller than that associated with a cost-free learning. Hence, in general, learning costs reduce the beneficial effect of plasticity but do not cancel it entirely.

Clearly, the fixed landscape, the one-to-one genotype/phenotype mapping and the specific phenotypic plasticity schemes examined in this study are a simplification of the dynamics that take place in natural systems. Living organisms incorporate a complex developmental process that may disturb the correlation between the phenotype and genotype spaces (Downing, 2004) and apply diverse and sophisticated plasticity methods. Moreover, the random walk model applied in this study restricts evolution to ± 1 increments, corresponding to small mutations. However, real mutations can come in a wide variety of increments, markedly influencing the evolutionary trajectory. Specifically, large mutations can help the evolutionary process to cross fitness valley barriers, reducing (though, most probably, not canceling altogether) the deleterious effect of the fitness landscape drawdown. The effect of such mutations on the evolutionary dynamics in innate vs. effective fitness landscapes is of much interest and may vary with the exact structure of the landscape. Unfortunately, characterizing the structure of fitness landscapes of biological systems and the dynamics of biological plasticity is still

an open question. We have thus focused on a simple mathematical model, with fully correlated landscapes, allowing us to explicitly construct and examine the resulting effective fitness function and to derive a rigorous, quantitative analysis of the expected convergence rate. However, the approach presented in this chapter and the mathematical analysis of first-passage times can be utilized to examine the effects of additional plasticity paradigms and landscapes structures, as long as the effective fitness landscapes can be evaluated.

Future research may be able to predict scenarios in which phenotypic plasticity will be favored by evolution. In particular, while the exact structure of the fitness landscape is usually hard to characterize, there are cases where multi-peaked landscapes are expected. For example, whenever the evolutionary process is required to optimize multiple objectives with a certain trade off function, multiple local optima usually exist (e.g., Oksanen and Lundberg, 1995). In such cases, phenotypic plasticity, being a mechanism that expedites the discovery rate of new optima, is valuable. We thus believe that increased phenotypic variation should be correlated with the existence and abundance of multiple optima.

An example of a biological experiment designed to directly examine the effect of learning on evolution has recently been presented (Mery and Kawecki, 2004). In this exciting study, populations of *Drosophila melanogaster* were exposed to various selection regimes concerning preference for oviposition substrate with and without the ability to use aversion learning. The results of this study showed that learning may evolve even in a stationary environment (in contrast to the common argument that learning should be favored only in a changing environment). However, examining whether learning ability affects the evolution of the innate component, they found that learning facilitated evolution in one direction of selection while hindering it in the other, leaving the controversy concerning the effect of learning on evolution unresolved.

Evidently, the dynamics governing the evolution of plastic individuals are still far from being completely characterized, much less understood. The variety of phenotypic plasticity mechanisms found in biological systems and their complexity render their study a challenging task. In this chapter, we have utilized a simple model to provide a quantitative analysis of these complex phenomena. While acknowledging its limitations, we believe that such a framework can serve as a theoretical basis for studying issues concerning the interplay between phenotypic plasticity and evolution.

One of the common and most exciting classes of phenotypic plasticity mechanisms is social learning, where learning takes place within a social situation and individuals learn by observing other members of the population. In the context of the interaction between learning and evolution, having other individuals in the population influence the behavior of a plastic individual may have profound effects on the evolutionary dynamics. For example, while it was shown that deterministic learning smooths the fitness landscape, clearly, the local nature of this learning scheme, wherein plastic individuals can perform only slight behavioral modifications in each learning iteration, bounds the effect of this learning process. In particular, the acquired behaviors are limited to those present in the local basin of attraction. Clearly, a less local learning strategy may further smooth the fitness landscape and further expedite the evolutionary process (as was demonstrate for stochastic learning). A social learning scheme, in which learners can examine the success (fitness) of other members of the population and can evaluate a wide range of behaviors, including those that considerably differ from their current behavior, has this exact property. To some extent, the experiments presented in Section 2.3.4 can be interpreted as modeling such a learning paradigm. Yet, a more direct model, integrating a parameter representing the population diversity and allowing a plastic individual to sample any improved behavior within this behavioral range, can provide valuable insights concerning the effect of social learning on evolution. We believe that such a learning scheme will further accelerate the rate of evolution in an analogous manner to that shown in Section 2.3.4. Furthermore, an improved model can also include a "population" of random-walkers and apply various selection operators to better encapsulate the range of behaviors present in an evolving population that can be observed by a social learner .

Acknowledging the significant influence of social learning on evolution, we turn in the rest of this dissertation to examine the effect of this class of learning mechanisms. However, while the social learning models discussed above can provide valuable insights concerning the interplay between learning and evolution (as was demonstrated in this chapter), such models still form a relatively abstract representation of that interplay, lacking many of the properties characterizing real instances of social learning. In the next chapter, we hence turn to examine one specific mechanism of social learning, namely, imitative learning. In terms of its potential effect on evolution, imitation presents far more complicated dynamics

as the learning process outcomes depend not only on the interaction between the individual and its environment but also on the state of other members of the evolving population. We resort to an Evolutionary Autonomous Agents (EAA) framework, facilitating an examination of the effect of a plausible model of imitative learning. Such an EAA model will also be utilized later (Chapter 4) to study the evolutionary origins of the mechanisms underlying learning by imitation.

Chapter 3

Enhancing Autonomous Agents Evolution with Learning by Imitation

Based on:

Elhanan Borenstein and Eytan Ruppin
Enhancing autonomous agents evolution with learning by imitation, *Journal of Artificial Intelligence and the Simulation of Behaviour (AISBJ)*, 1(4), 335-347, 2003.

In the previous chapter we have shown that phenotypic plasticity can accelerate the evolutionary process and have identified the origins of this beneficial effect. To allow a rigorous analysis of the effect of plasticity we have confined our study to plasticity schemes in which the resulting effective fitness landscape can be explicitly constructed. However, the various plasticity and learning mechanisms found in nature are often significantly more complex, preventing a full mathematical analysis. One such exciting learning mechanism is learning by imitation.

The main motivation for studying the effect of learning by imitation on evolution is twofold (see also Section 1.2.2). First, imitation is an effective and robust way to learn new traits by utilizing the knowledge already possessed by others. Although the existence of true imitative behavior in the animal kingdom is still debated, social learning can be found in a variety of species providing clear benefits over other forms of learning (Kawamura, 1963; Whiten and Ham, 1992; Zentall, 2001). Second, in the context of evolutionary computation and

autonomous agents' evolution, the inherent diversity of the agents' population makes learning by imitation a naturally applicable mechanism. Moreover, while other forms of supervised learning may not be a viable option in many agent environments due the the lack of training data (or Oracles), learning by imitation is almost always a valid option, using other members of the population as teachers.

In contrast to simple phenotypic plasticity mechanisms, an imitative learning process depends not only on the interaction between the individual and its environment but also on the state of other members of the evolving population, making the interaction between this learning scheme and evolution far more complex. In this chapter, we hence focus on the effect of learning by imitation on the evolutionary process, utilizing a framework of autonomous agents. We present a set of experiments, where lifetime learning by imitation was used to adapt individuals that also go through an evolutionary process. The results are compared with those of a simple evolutionary process, where no lifetime learning is employed, and with those of an evolutionary process that employs a conventional (supervised) mechanism of learning.

The remainder of this chapter is organized as follows. We begin in Section 3.1 with a brief overview of previous work and the concept of combining imitative learning with evolution. In Section 3.2 we present the *imitation enhanced evolution (IEE)* model in detail. To validate the effectiveness of our model we introduce in Section 3.3 a set of benchmark tasks and present the experimental results in Section 3.4. The chapter concludes with a short discussion concerning the results obtained. The experiments and results presented in this chapter have been published in Borenstein and Ruppin (2003a) and Borenstein and Ruppin (2003b).

3.1 Combining Learning by Imitation and Evolution

As discussed in the previous chapter, various simulation studies of the interaction between lifetime learning and evolution (Hinton and Nowlan, 1987; Nolfi et al., 1994; Nolfi and Parisi, 1997; Floreano and Mondada, 1996) have shown that learning can be utilized to guide and enhance the evolutionary process. While the mathematical analysis presented in Chapter 2 provides a rigorous account for this effect, an intuitive explanation for the beneficial effect of learning on evolution can also be introduced. As illustrated in Figure 3.1, learning can be conceived

as a simple mechanism that reveals the innate potential of each individual in the population. One may consider lifetime adaptation as a local search process that can enhance the global search (evolution) by determining which configurations lie in the vicinity of the global optimum solution and are thus worthwhile retaining in the population (as they have a better chance to produce successful offspring).

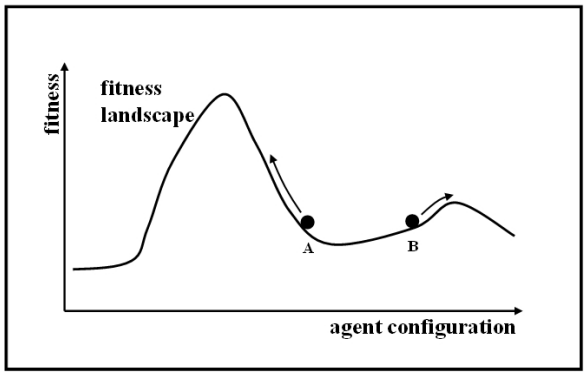


Figure 3.1: *An illustration of the effect that lifetime adaptation may have on the genetic evolutionary process.* Both agents start with the same innate fitness value (indicated by the black dots). Applying lifetime adaptation (illustrated as a simple hill climbing process) will result in the selection of agent A which is closer to the optimal solution. Inspired by Nolfi and Floreano (1999)

The studies cited above have employed various sources of training data for learning ranging from external oracles, through regularities in the environment, to "self-generated" teaching data. There is, however, an additional source of training data; one which is naturally available within the evolutionary paradigm - the knowledge possessed by other members of the population. Such knowledge can be harnessed to improve the evolutionary process in the form of *learning by imitation*. Extending these studies further, we hence wish to explore learning by imitation as an alternative to conventional supervised learning and to apply it as a tool to enhance evolution. We will label this framework as *imitation enhanced evolution (IEE)*. The hypothesis underlying the study presented in this chapter is that learning by imitation, may be sufficient to reveal the innate potential of the population members.

Learning by imitation has already been applied by researchers in the fields of artificial intelligence and robotics in various experiments. Hayes and Demiris (1994) presented a model of imitative learning to develop a robot controller. Billard and Dautenhahn (1999) studied the benefits of social interactions and imi-

tative behavior for grounding and use of communication in autonomous robotic agents. For an up-to-date introduction to work on imitation in both animals and artifacts see the cross-disciplinary collection (Dautenhahn and Nehaniv, 2002b). Furthermore, various frameworks that study the interaction between cultural transmission and evolution have already been well established (e.g. Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981; Laland, 1992). Gene-culture coevolution accounts for many adaptive traits (Feldman and Laland, 1996). Studies and simulations of the evolution of language (Ackley and Littman, 1994; Kirby and Hurford, 1997; Arbib, 2002) assume, by definition, some sort of cultural transmission.

It is important to realize though, that in contradistinction to these studies, the framework presented in this chapter does not employ cultural evolution. In fact, we preclude culture from evolving in the first place. Following in the footsteps of the studies of the interaction between learning and evolution discussed above, we avoid any direct form of acquired-knowledge transfer between generations either genetically or culturally. We hence work in a strict Darwinian framework, where lifetime adaptations are not directly inherited (although, they may be genetically assimilated through the Baldwin effect, 1896) and may affect the evolutionary process only by changing the individual's fitness, and thus the number of its offspring. In terms of cultural transmission (see Boyd and Richerson, 1985, for a detailed definition), we allow *horizontal* transmission alone (where individuals of the same generation imitate each other) and exclude any form of *vertical* transmission (where members of the current generation transmit their knowledge to members of the next generation). Numerous field studies suggest that at least in nonhuman societies, horizontal transmission is far more common than vertical transmission (Laland, 1992). Furthermore, to prevent any form of cultural evolution from taking place, within each generation, only innate behaviors are imitated; that is, we prevent behaviors acquired by imitation from being imitated again by another member (see also the discussion in Section 3.5).

Within this framework, we focus on the effects that imitation may have on the genetic evolutionary process, starting with the most basic question: Can imitation enhance the evolution of autonomous agents (in the absence of vertical transmission), in an analogous manner to the results previously shown for other forms of learning, and how? It should be noted that the contribution of horizontal transmission to evolution is not obvious; while in late stages of the

evolutionary process the best individuals in the population may already possess sufficient knowledge to approximate a successful teaching process, in early stages of the process it may be the case of "the blind leading the blind", resulting in a decrease of the population's average fitness.

A simple model that fits this framework has been studied before by Best (1999). He demonstrated an extension of the computational model presented in Hinton and Nowlan (1987), introducing social learning (namely imitation) as an additional adaptive mechanism. The reported results exemplify how horizontal cultural transmission can guide and accelerate the evolutionary process in this simplified model. Best has also demonstrated how social learning may be superior to conventional learning and yield faster convergence of the evolutionary process.

Best's model, however, has several limitations. The evolutionary fitness function (which is the one used in Hinton and Nowlan, 1987) represents a worst-case scenario where only the exact solution has a positive fitness value. There is no probable path that a pure evolutionary search can take to discover this solution. As we have demonstrated in Chapter 2, the structure of the fitness landscape plays a crucial role in the interplay between learning and evolution. It is only in the biologically plausible case of multipeaked landscapes that the true effect of learning can be revealed. Additionally, in Best's model, there is no distinction between genotypes and phenotypes and thus no real *phenotypic* adaptation process. Imitation is carried out simply by copying certain genes from the teacher's genome to the student. A complex mapping from genotypes to phenotypes is one of the important factors determining the characteristics of the fitness landscape and consequently, the effect of learning on evolution. We hence wish to generalize this framework and study the effects of learning by imitation in a more realistic scenario of *autonomous agents* evolution (see Ruppin, 2002, for a general review). A framework based on such evolutionary autonomous agents allows for a simple test-bed for studying learning by imitation and provides a natural mechanism for genotype-phenotype mapping that can produce nontrivial fitness landscape structures.

The definition of imitation in the literature varies considerably (Billard and Dautenhahn, 1999), but for the purpose of this study we use imitation (or learning by imitation) in the sense of having an individual (student) being able to match its behavior to that of a demonstrator (teacher). In particular, using autonomous agents to model the population members, this form of imitation is implemented

by using the teacher's output for each sensory input as the target output in a back-propagation training algorithm. In this study we focus on the simple case where the learning (imitation) task is similar to the evolutionary task. This case most probably does not closely represent the imitation processes found in nature. Lifetime adaptation in humans and other cultural organisms operates on high-level traits which are not coded directly in their genome. However, we believe that this simple scenario can provide valuable insights into the roots of imitative behavior. We further discuss this topic in Section 3.5.

3.2 The Model and Experimental Setup

A haploid population of agents evolve to solve various tasks. Each agent's neuro-controllers is a simple feed-forward (FF) neural network (Hertz et al., 1991). The initial weights of the network synapses are coded directly into the agent's genome (the network topology is static throughout the process). The initial population is composed of 100 individuals, each assigned randomly selected connection weights from the interval [-1,1]. The *innate* fitness of each individual is determined by its ability to solve the specific task upon birth. Within the pure evolutionary process, the innate fitness will determine the reproductive probability of this individual. Each new generation is created by randomly selecting the best agents from the previous generation according to their innate fitness, and allowing them to reproduce (Mitchell, 1998). During reproduction, 10% of the weights are mutated by adding a randomly selected value from the interval [-0.35,0.35]. The genomes of the best 20 individuals are copied to the next generation without mutation.

When conventional supervised learning is applicable (i.e., an explicit oracle can be found) we also examined the effect of supervised learning on the evolutionary process. Each individual in the population goes through a lifetime learning phase where the agent employs a back-propagation algorithm (Hertz et al., 1991), using the explicit oracle as a teacher. Its fitness is then reevaluated to determine its *acquired* fitness (i.e., its fitness level after learning takes place). In order to simulate the delay in fitness acquisition associated with acquired knowledge, we use the average of the innate and acquired fitness values as the agent's *final* fitness value. This delay is a simple manifestation of the indirect costs of learning (in comparison to innate, genetically encoded, traits), stemming from the non-optimal behavior displayed by the individual until it acquires a successful

behavior by learning. This averaged fitness value is then used to select the agents that will produce the next generation.

In the IEE paradigm, agents do not use conventional supervised learning, but rather employ learning by imitation. In every new generation of agents, created by the evolutionary process, each agent in the population selects one of the other members of the population as an imitation model (teacher). Teachers are selected stochastically, where the probability of selecting a certain agent as a teacher is proportional to its innate fitness value (i.e., its initial fitness levels before learning takes place). The agent employs a back-propagation algorithm, using the teacher's output for each input pattern as the target output, mimicking a supervised learning mode. The imitation phase in each generation can be conceived as happening simultaneously for all agents, preventing behaviors acquired by imitation from being imitated. Only the innate behavior of the teacher is imitated by the student. The acquired fitness and final fitness are evaluated in the same method that was described in the case of conventional learning.

As stated above, acquired knowledge does not percolate across generations. Each time a new generation is produced, all lifetime adaptations possessed by the members of the previous generation are lost. Newborn agents inherit only the genome of their parents, which does not encode the acquired network adaptations that took place during the parent's lifetime. Successful individuals that were copied from the previous generation also go through a new genotype-to-phenotype ontogenetic development process and thus lose all adaptations acquired during the previous generation.

To summarize, learning by imitation in a population of evolving agents (IEE) works as follows:

---

1. *Create the initial population. Assign the network weights of each individual with randomly selected values.*

2. *Repeat:*

   (a) *For each individual in the population:*

      i. *Evaluate the innate fitness F_i.*

   (b) *For each individual S in the population:*

      i. *Set S to be the student.*

ii. *Select a teacher T from the population. The probability of selecting a certain individual as a teacher is proportional to its innate fitness value F_i.*

iii. *Train S with back-propagation algorithm. Use the output of T as the desired output (when computing the output of T, use the innate configuration of T).*

iv. *Evaluate the acquired fitness F_a of S.*

(c) *For each individual in the population:*

i. *Evaluate the final fitness $F_f = \frac{F_i + F_a}{2}$.*

(d) *Create the next generation by selecting the best individuals according to F_f and allow them to reproduce as described above.*

3.3 The Benchmark Tasks

The model described in the previous section was tested on three different tasks. The first two are standard classification benchmark problems. The third is an agent-related task used in previous studies of the interaction between learning and evolution.

3.3.1 The Parity Problem

The agents evolved to solve the five bit parity problem. A network topology of 5-6-2-1 was used (i.e., 5 input neurons, two hidden layers, the first with 6 neurons and the second with 2, and 1 output neuron), with an additional threshold unit in each layer. All 32 possible input patterns were used both for evaluating the network performance and for training.

3.3.2 The Classification Problem

A simple two-dimensional geometrical classification problem was used in this task. The network receives as input a point from the unit square and should determine whether it falls within the boundaries of a predefined triangle. A network topology of 2-5-1 was used (with an additional threshold unit in each layer). The test set and training set consisted of 100 points randomly selected from the unit square.

3.3.3 The Foraging Task

The task in this simulation is similar to the one described by Nolfi et al. (1994). An agent is placed on a two-dimensional grid-world (Figure 3.2). A number of food objects are randomly distributed in the environment. As its sensory input the agent receives the angle (relative to its current orientation) and distance to the nearest food object. The agent's output determines one of four possible actions: turn 90 degrees left, turn 90 degrees right, move forward one cell, or do nothing (stay). If the agent encounters a food object while navigating the environment, it consumes the food object. The agent's fitness is the number of food objects that were consumed during its lifetime. Each agent lives for 100 time steps in a 30x30 cells world which initially contains 30 food objects. A network topology of 2-6-2 was used (with an additional threshold unit in each layer).

In this task, unlike the previous ones, there is no explicit oracle we can use to train the agent. Nolfi et al. (1994) used available data to train the agent on the task of predicting the next sensory input, which differs, but is in some sense still "correlated" with that of finding food (the evolutionary task). In our model, we can still use the same mechanism of learning by imitation to train the agent on the original evolutionary task, using the best individuals in the population as teachers.

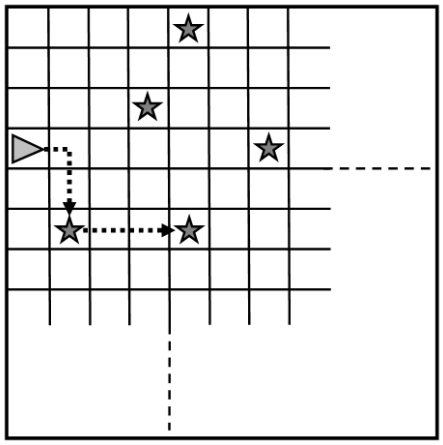


Figure 3.2: *The foraging task*: The agent (triangle) navigates in a 2D grid-world. Food objects (stars) are randomly distributed in the world. The agent can turn 90 degrees left, turn 90 degrees right, move one cell forward, or stay. Each time the agent encounters a food object, it consumes the food object and gains one fitness unit. Inspired by Nolfi and Floreano (1999)

There are several strategies we can apply to determine which sensory input patterns should be used for training. Randomly selecting arbitrary input patterns, as we did in previous tasks, is not a suitable strategy here as the real input distribution that an agent encounters while navigating the environment may differ considerably from a uniform distribution. However, two behaviorally motivated strategies may be considered: a *query* model and an *observational* model. In the query model, the student agent navigates in the environment and for each sensory input pattern it encounters, the student queries the teacher to obtain the teacher's output for this pattern. The teacher's output is than used as the target output in back-propagation training of that pattern. In the observational model, the student "observes" the teacher agent as the teacher navigates in the environment and uses the sensory input patterns encountered by the teacher as training patterns (again, using the teacher's output for the back-propagation algorithm). Using this model we can further limit the observed patterns to those which occur during time steps that precede the event of finding food. This constraint will allow the student to imitate only useful behavioral patterns. We will label this strategy as *reinforced agent imitation (RAIL)*.

3.4 Results and Analysis

We first studied IEE in the two classification tasks described in the previous Section, where conventional supervised learning can still be applied. In these tasks we were able to compare the effects that both lifetime adaptation mechanisms (i.e., learning and imitation) have on the evolutionary process and to specifically examine what are the effects of using imitative data rather then real data. The results clearly validate that the IEE model consistently yields an improved evolutionary process. The innate fitness of the best individuals in populations generated by applying learning by imitation is significantly higher than that produced by standard evolution.

Figure 3.3 illustrates the innate performances of the best agent as a function of generation, in populations evolved to solve the triangle classification problem (Section 3.3.2). To evaluate the agent's classification accuracy we use the Mean-Square Error (MSE) measure to calculate the distance between the network predicted classification and the true classification, averaged over all the patterns in the test set. Fitness is defined as $(1-Error)$. The results of a simple evolutionary

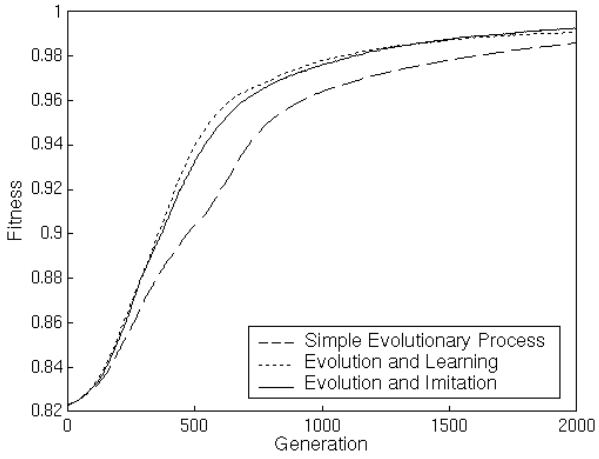


Figure 3.3: *The triangle classification task*: the innate fitness of the best individual in the population as a function of generation.

process (dashed line) and of an evolutionary process that employs conventional supervised learning (dotted line) are compared with those of an evolutionary process that employs learning by imitation (solid line). Each curve represents the average result of 4 different simulation runs with different, randomly assigned, initial connection weights (the number of repeated simulation runs is limited by the extensive computation time required in each generation). The results presented in Figure 3.3 demonstrate how applying either of the learning paradigms yields better performing agents than those generated by a simple evolutionary process. Furthermore, applying learning by imitation produces practically the same improvement throughout the process as does conventional supervised learning.

When facing the 5-bit parity task, the effect of applying lifetime adaptation is even more surprising. Figure 3.4 illustrates the innate performances of the best agent as a function of generation, in populations evolved to solve the 5-bit parity problem. Each curve represents the average result of 10 different simulation runs with different, randomly assigned, initial connection weights. While simulations applying the IEE model still outperform the simple evolutionary process, using conventional supervised learning actually results with a significant decrease in performances. The problematic nature of this specific task may account for these poor results. The parity problem, although often used as a benchmark, is considered to be a difficult and atypical classification problem (Fahlman, 1989). Learning algorithms facing this task tend to get trapped in local minima. However, learning from an imperfect teacher, as is the case in learning by imitation,

induces a certain level of noise into the learning process and may thus help to prevent the process from getting stuck.

Evidently, learning by imitation has a similar (and in some cases, superior) effect on the evolutionary process to the one that was previously shown for conventional supervised learning. The knowledge possessed by the best members of the population can be used as alternative training data for other members, even in the early stages of the evolutionary process. We then turned to use IEE to enhance evolution where explicit training data is not available. This is the case in the foraging task described in Section 3.3.3.

Figure 3.5 illustrates the results of the simulations in which the agents faced the foraging task. The average innate fitness of the population in a simple evolutionary process is compared with the average innate fitness of populations that applied learning by imitation. The agents in this simulation employed an observational model and used the *RAIL* strategy of imitation. Fitness is measured as the number of food objects an agent consumes during its lifetime. Each curve represents the average result of 10 different simulation runs with different, randomly assigned, initial connection weights. As can be seen in Figure 3.5, autonomous agents produced by our model demonstrate better performances than those generated by the simple evolutionary process; that is, their innate capacity to find food in the environment is superior.

We also examined the effect of employing different *adaptation levels*. In our experimental setup, the adaptation level is implemented simply as the number of

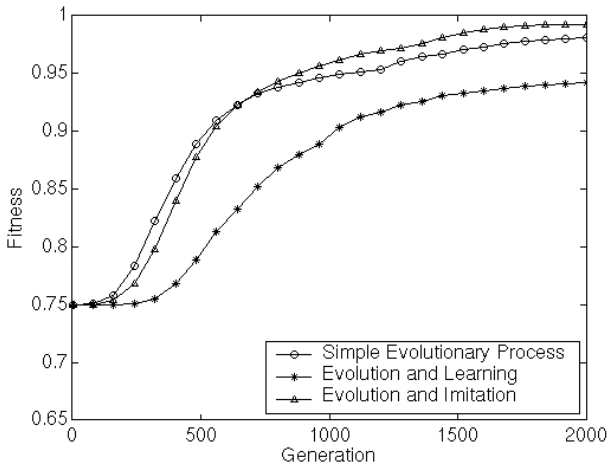


Figure 3.4: *The 5-bit parity task*: the innate fitness of the best individual in the population as a function of generation.

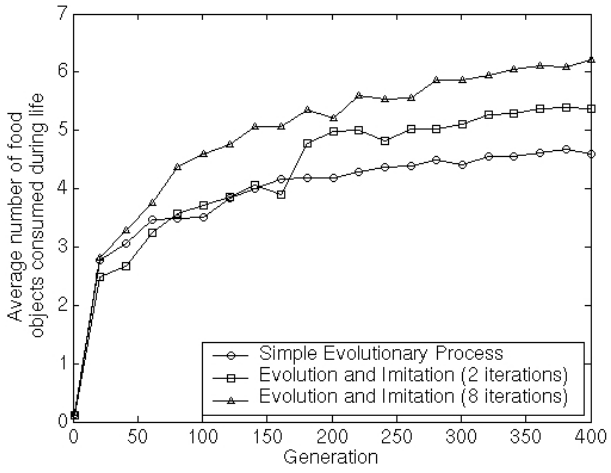


Figure 3.5: *The foraging task: the average innate fitness of the population as a function of generation.* The results of a simple evolutionary process are compared with those of simulations that employed lifetime imitation with two distinct adaptation forces (2 and 8 learning iterations). Imitating agents employed an observational model, using the *RAIL* strategy.

learning (back-propagation) iterations we apply in each lifetime adaptation phase. The results illustrated in Figure 3.5 also demonstrate that a higher adaptation level (i.e., a higher number of iterations in each imitation phase) further improves the performance of the resulting agents. This effect coincides with an analogous effect reported by Best (1999) where higher transmission force resulted with faster convergence of the evolutionary process.

To further explore the effects of lifetime imitation on evolution, we examined the improvement in fitness during lifetime as a function of generation. The improvement can be evaluated by calculating the difference between the acquired fitness and the innate fitness (i.e., $F_a - F_i$) in every generation. The results illustrated in Figure 3.6A clearly demonstrate that in very early stages of the evolutionary process, the best agents in the population already possess enough knowledge to improve the fitness of agents that imitate them. In fact, the contribution of imitative learning decreases as the evolutionary process proceeds, probably due to population convergence to high performance solutions.

An additional observation on the interaction between lifetime adaptation and evolution can be obtained from examining the diversity of the population throughout the evolutionary process. The average genome variance of the population, i.e., the variance among the population members, in the value of each gene (encoding a certain network weight) averaged over all genes, can serve as a measure of the

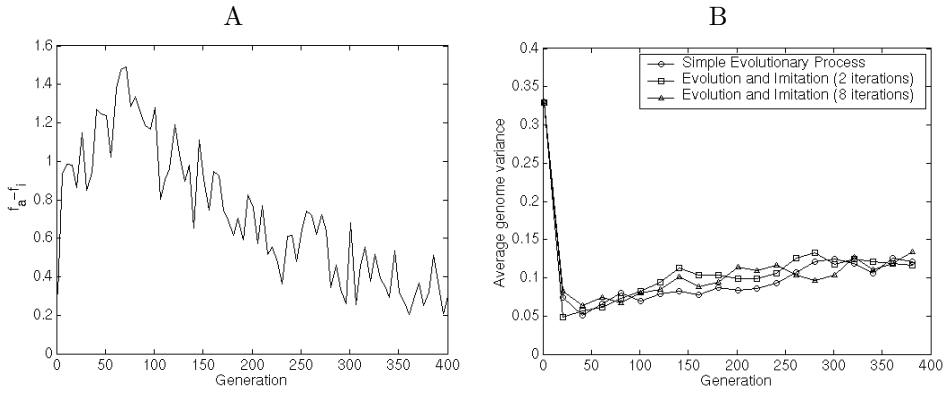


Figure 3.6: *The foraging task: the improvement of the population average fitness gained by lifetime imitation (***A***) and average genome variance (***B***) as a function of generation.* Populations that employ lifetime adaptation, maintain a higher diversity throughout the evolutionary process.

population's diversity. As demonstrated in Figure 3.6B, during the first few generations, the population's initial diversity decreases rapidly due to the selection pressure of the evolutionary process. However, throughout most of the following generations, the diversity found in populations subject to lifetime adaptation by imitation is higher than the diversity of populations undergoing a simple evolutionary process. Allowing members of the population to improve their fitness through lifetime adaptation before natural selection takes place facilitates the survival of suboptimal individuals and helps to maintain a diversified population. This feature can partly account for the benefit gained by applying lifetime adaptation to agents evolution.

3.5 Discussion

This chapter focuses on the effects of imitation on the evolution of agents in the absence of cultural evolution. We show that introducing the adaptive mechanism of lifetime learning by imitation, utilizing the knowledge possessed by members of the population, can significantly improve the evolutionary processes and result in better performing agents. Our IEE model proves to be a powerful tool that can successfully enhance the evolution of autonomous agents attempting to successfully solve various tasks. This paradigm is particulary useful in evolutionary simulations of autonomous agents, where conventional supervised learning is not possible.

The imitation paradigm presented in this chapter assumes that the agents can estimate the fitness of their peers (i.e., more successful agents are larger and look healthier, etc.). More specifically, the RAIL strategy, where agents imitate only successful behavior, assumes that agents can detect significant changes in the fitness of their peers during their lifetime or identify specific activities that may contribute to their fitness. In principle, the model presented in Section 3.2 can provide a framework to explore ways in which such assumptions can be relaxed. Coding imitative behavior attributes (such as the imitation model selection scheme, imitation strategy, imitation period, etc.) into the genome might result in the spontaneous emergence of intriguing imitative behavior patterns in a population of agents.

The model can also be extended to study the incentive that should be provided to an agent to make it assume the role of a teacher. Teaching, or even allowing someone else to imitate one's actions is, by definition, an altruistic behavior, and might have various costs associated with it. There may hence be specific conditions which can lead to the emergence of active teaching even in the presence of a fitness penalty for such a behavior. Such favorable teaching conditions may arise when the fitness associated with various actions is correlated with the frequency of these actions in the population (see also Boyd and Richerson, 1985, for a discussion of frequency-dependent bias). A good example of this case can be found in the emergence of normative behaviors (Axelrod, 1986; Flentge et al., 2001).

Addressing the question concerning the effect of imitation on evolution (rather than the interaction between culture and evolution), the IEE model presented here entails a relatively simple form of cultural transmission, confined to a horizontal transmission alone. Although horizontal transmission may be more common than vertical transmission (Laland, 1992), in general, cultural evolution is a more complex process, involving both horizontal and vertical transmission (as well as other forms of cultural transmission), and a parallel evolution of multiple interacting traits. Such cultural evolution processes may produce dramatically different dynamics (see, for example, Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981; Laland, 1992) and will be explored in Chapter 5.

Moreover, to allow a direct comparison with previous studies focusing on the interaction between learning and evolution, we also avoid any direct form of acquired-knowledge transfer even within the same generation, and limit the

imitative process to imitation of innate behaviors. Clearly, imitating behaviors that are the result of ontogenetic imitation by other individuals can have significant consequences and produce markedly dissimilar dynamics. For example, in the extreme case where ongoing imitation of acquired behaviors is not limited, one can expect the entire population to eventually converge to the same behavior. Although the underlying genetic diversity will still be maintain, selection pressures will be totally suppressed, bringing the evolutionary process into a dead-end. In such a scenario, imitation could hinder the evolutionary process rather than enhance it. Even if such an extreme case is not assumed, we suspect that imitating acquired behaviors may push individuals further and further away from their innate configuration, masking their genetic potential and consequently slow the evolutionary process.

Most importantly, as this chapter focuses on the effect of imitative learning on evolution, in our model the agents' ability and incentive to imitate is assumed to be instinctive. Quoting Billard and Dautenhahn (1999), "our experiments address learning by imitation instead of learning to imitate". However, the mechanisms underlying imitative learning are themselves the product of an evolutionary process. It is thus intriguing to also examine the evolutionary origins and emergence of the neuronal devices that give rise to imitative behavior and to identify the fundamental characteristics of this evolutionary trajectory. In the following chapter, we specifically address this question, presenting a novel framework for studying the evolutionary origins of imitative behavior.

Chapter 4

The Emergence of Imitation and Mirror Neurons in Adaptive Agents

Based on:

Elhanan Borenstein and Eytan Ruppin
The evolution of imitation and mirror neurons in adaptive agents, *Cognitive Systems Research (special issue on Epigenetic Robotics)*, 6(3), 229-242, 2005.

Elhanan Borenstein and Eytan Ruppin
The evolutionary link between mirror neurons and imitation: An evolutionary adaptive agents model, *Behavioral and Brain Sciences*, 28:2, 127-128, 2005.

Elhanan Borenstein and Eytan Ruppin
Evolving imitating agents and the emergence of a neural mirror system, *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*, 146-151, 2004.

The interplay between any learning mechanism and evolution is clearly bidirectional: As we have shown in Chapter 2 and Chapter 3, the presence of lifetime adaptation (namely, phenotypic plasticity and imitation) can dramatically affect the dynamics of the evolutionary process, yielding, in some cases, an acceler-

ated and enhanced evolution. However, it should be noted that these adaptation mechanisms are themselves the product of an evolutionary process and may hence be properly understood only in light of their evolutionary history.

The above statement is especially true for imitative learning. In Chapter 3 we have examined the effect of imitative learning on the evolution of autonomous agents, assuming that the capacity to imitate is present. However, imitation is a highly complex cognitive process, involving vision, perception, representation, memory and motor control, making the study of its evolutionary origins an interesting task. In this chapter, we thus turn to examine the evolution of the mechanism underlying imitative behavior.

The remainder of this chapter is organized as follows: We begin in Section 4.1 and Section 4.2 with a brief overview of the mirror neurons phenomena and discuss the appropriate method for studying the evolutionary origins of the neuronal mechanism underlying imitative behavior. We then set out to pursue two objectives: We first, in Section 4.3, present a novel experimental framework for evolving imitative learning in evolutionary adaptive autonomous agents (Ruppin, 2002; Floreano and Urzelai, 2000). We demonstrate the evolution of imitating agents that comprise a simple mechanism of imitative behavior that was not explicitly engineered into the agents. We then turn, in Section 4.4, to systematically analyze the structure and dynamics of the resulting neurocontrollers. This analysis reveals neural devices analogous to those found in biological systems, including clear examples of internal coupling between observed and executed actions. Further analysis of the network adaptation dynamics reveals a hybrid mechanism, combining an innate perceptual-motor coupling with acquired context-action associations. The chapter concludes with a discussion of the implications of our findings for imitation theory. The study presented in this chapter has been published in Borenstein and Ruppin (2004), Borenstein and Ruppin (2005a) and Borenstein and Ruppin (2005b).

4.1 Imitation and Mirror Neurons

The past twenty years have seen a renewed interest in imitation in various fields of research (Prinz and Meltzoff, 2002) such as developmental psychology (Meltzoff, 1996), experimental studies of adult social cognition (Bargh, 1997), and most relevant to our work, neurophysiology and neuropsychology (Rizzolatti et al.,

1996, 2002). Research in this last field has led to the exciting discovery of *mirror neurons*. These neurons were originally found in the ventral premotor cortex (area F5) in monkeys, an area which is characterized by neurons that code goal-related motor acts (e.g. hand or mouth grasping). Some of the neurons in this area, which have been termed *mirror neurons*, discharge both when the monkey performs an action and when it observes another individual making a similar action (Gallese et al., 1996; Rizzolatti et al., 2002). Most mirror neurons exhibit a marked similarity in their response to action observation and execution, and in some cases this similarity is extremely strict (Rizzolatti et al., 2001). An analogous mechanism, whereby cortical motor regions are activated during movement observations was also demonstrated in humans using TMS (Fadiga et al., 1995), MEG (Hari et al., 1998), EEG (Cochin et al., 1998) and fMRI (Iacoboni et al., 1999; Buccino et al., 2001). Mirror neurons are thus the first identified neural mechanism that demonstrates a direct matching between the visual perception of an action and its execution. The ability to match the actions of self and other may have a functional role in fundamental cognitive processes, such as understanding the actions of others, language and mind reading (Rizzolatti et al., 2001). In particular, imitation of motor skills requires the capacity to match the visual perception of a demonstrator's action to the execution of a motor command. The neural mirror system, demonstrating such an internal correlation between the representations of perceptual and motor functionalities, may form one of the underlying mechanisms of imitative ability.

4.2 Evolving Imitating Agents: Emerged vs. Engineered Approach

4.2.1 Context-Based Imitation

Learning by imitation, like any cognitive process, must be considered an intrinsically embodied process, wherein the interaction between the neural system, the body and the environment cannot be ignored (Keijzer, 2002; Dautenhahn and Nehaniv, 2002a). In particular, every action, either observed or performed, occurs within a certain *context*. A context can represent the time or place in which the action is made, various properties of the environment, the state of the individual performing the action or the social interaction partners (see, for example, Dautenhahn, 1995). Clearly, there is no sense in learning a novel behavior by

imitating another's actions if you do not know the context in which these actions are made – a certain action can be extremely beneficial in one context, but have no effect (or even be deleterious) in a different context. Discussing an agent-based perspective on imitation, Dautenhahn and Nehaniv (2002a) consider the problem of imitating the right behavior in the appropriate context, i.e., "when to imitate", as one of the five central questions ("Big Five") in designing experiments and research on imitation. We hence use the term *context-based imitation* in the sense of being able to reproduce another's observed action whenever the context in which the action was originally observed, recurs.[1] For example, an infant observing his parents may learn by imitation to pick up the phone (*action*) whenever the phone is ringing (*context*).

Context-based imitation can thus be conceived as constructing a set of associations from contexts to actions, based on observations of a demonstrator performing different actions within various contexts. These associations should comply with those that govern the demonstrator's behavior, and should be learned (memorized) so that each context stimulates the production of the proper motor action even when the demonstrator is no longer visible. It should be noted however, that "*action*" is an abstract notion, and in reality, an imitating individual (agent) should also be capable of matching a *visual perception* of the demonstrator's action to the corresponding *motor command* that activates this action.[2] The key objective of this study is to gain a comprehensive understanding of the mechanisms that govern such context-based imitative learning and to examine the nature of the associations between visual perception, motor control and contexts that are being formed in the process.

4.2.2 Studying the Origin of Imitation

Imitation is an effective and robust way to learn new traits by utilizing the knowledge already possessed by others and it has already been applied by researchers in the fields of artificial intelligence and robotics (see Section 3.1). Furthermore, some researchers, motivated by the recent discovery of a neural mirror

[1] Animal behavior and human psychology literature introduces a wide range of definitions of imitation, focusing on what can constitute true imitation vs. other forms of social learning (Zentall, 2001; Nehaniv and Dautenhahn, 2002). Our definition addresses the importance of the observed action's *context* for a successful behavior.

[2] In this study we focus on visually based imitation. However, it should be noted that other forms of imitation, such as vocal imitation, need not involve visual modality (see, for example, Nehaniv and Dautenhahn, 2002; Herman, 2002).

system, have implemented various models for imitative learning, employing neurophysiologically inspired mechanisms. Billard (2000) presented a model of a biologically inspired connectionist architecture for learning motor skills by imitation. The architecture was validated through a mechanical simulation of two humanoid avatars, learning several types of movements sequences. Demiris and Hayes (2002) and Demiris and Johnson (2003) developed a mirror-neuron based computational architecture of imitation inspired by Meltzoff's Active Intermodal Matching mechanism (Meltzoff and Moore, 1997) and combined it with an "active" distributed imitation architecture. They have demonstrated that this dual-route architecture is capable of imitating and acquiring a variety of movements including unknown, partially known, and fully known sequences of movements. Oztop and Arbib (2002), focusing on the grasp-related mirror system, argued that mirror neurons first evolved to provide visual feedback on one's own "hand-state" and were later generalized to understanding the actions of others. They have conducted a range of simulation experiments, based on a schema design implementation of that system, providing both a high-level view of the mirror system and interesting predictions for future neurophysiological testing. Other researchers (Marom et al., 2002; Kozima et al., 2002) claimed that the mirror system structure can be acquired during life through interaction with the physical or social environment and demonstrated models whereby perceptual and motor associations are built up from experience during a learning phase.

The studies cited above, however, assume that the agents' basic ability and motivation to imitate are innate, explicitly introducing the underlying functionality, structure or dynamics of the imitation mechanism into the experimental system. **In contrast to this engineering-based approach, we wish to study the neuronal mechanisms and processes underlying imitation from an evolutionary standpoint, and to demonstrate how imitative learning *per se* can evolve and prevail**. Evolutionary autonomous agents form an intuitively appealing approach for modeling and studying the evolution of biological neural mechanisms (Ruppin, 2002). Using a simulated environment, wherein agents *evolve* to perform a simple imitative task, facilitates a thorough examination of the resulting mechanism in "ideal conditions": Full control of the environment and experimental setup, and perfect knowledge of the agents' behavior and neural dynamics. Clearly, acknowledging the evolutionary origins of imitation and examining the emerging (rather than engineered) device can shed

light on the common *fundamental* principles that give rise to imitative behavior. It is important to note, however, that our key goal in this model is *not* to simulate the neural mechanism that underlies imitative behavior in the human or primate brain nor to incorporate the full range of social skills required for imitative learning (e.g. extraction of the context from the environment or coping with a different embodiment). The model described in this chapter is clearly a simplified conceptual model and does not presume to encapsulate many of the well established biological and neuronal data on imitation. Rather, the aim of such an evolutionary autonomous agents model is to examine *generic* and *universal* properties of complex living systems (the *"life as it could be"* paradigm, Langton, 1988, 1995). *The key point in this study is thus to examine the emerging characteristics of the mechanism evolved to support imitation in a system where no constraints on the underlying mechanisms or representations were explicitly encoded.*

4.3 The Experimental Setup

4.3.1 The Environment

The agents in our simulation inhabit a world that can be in one of several *world states* $\{s_1, s_2, \ldots, s_n\}$. In each time step, the world state is randomly selected from $\{s_1, s_2, \ldots, s_n\}$ with a uniform distribution. These states can represent, for example, the presence of certain food items or the size of an observed object and hence form the *context* in which actions are observed and performed. The world state, however, is not visible in every time step and is seen (i.e. included in the agent's sensory input) only in 60% of the time steps. An additional set, $\{a_1, a_2, \ldots, a_m\}$, represents the repertoire of motor *actions* that can be performed by the agent or by the demonstrator. A *state-action injective mapping* is also defined, assigning a certain action as the proper action for each world state s_i. Within the simulations described below, both n and m are set to 4, allowing $4! = 24$ different state-action mappings. Regularly performing the proper action assigned to the current state of the world is deemed a successful behavior and confers a positive fitness. Similarly, when the world state is not visible, a successful agent should not perform any action. It is assumed that the environment is also inhabited by a demonstrator (teacher), successfully performing the proper action in each time step. The demonstrator's action is visible (i.e. included in

the agent's sensory input) only in 20% of the time steps. The partial visibility of the world state and demonstrator ensures that during the agent's life it will encounter both scenarios wherein the demonstrator is not visible, forcing the agent to "memorize" the proper state-action mapping, and scenarios wherein the world state is not visible, in which a successful agent should "observe" the demonstrator's action but not perform any action. The specific visibility values defined above have no significant effect on the resulting agent, but rather provide a good blend of the various visibility scenarios during the agents' life, facilitating the examination of the agents' neurocontroller in these scenarios. **Furthermore, the above mapping, from world states to actions, is randomly selected anew in the beginning of each agent's run in the world**. The motivation for this state-action mapping shuffle is twofold. First, it prevents such a mapping from becoming genetically determined. To demonstrate a successful behavior, agents must *learn* the proper mapping by observing the demonstrator, promoting an imitation based mechanism to evolve. Second, it represents a scenario of a changing environment, wherein novel world states appear over time (new food sources, other species, etc.), making prior state-action mappings obsolete.

4.3.2 The Agent

Figure 4.1 illustrates the structure of the agent's sensorimotor system and neurocontroller. The agent's sensory input in each time step comprises 8 binary values, including the current world state (if visible) and a 4-cell retinal "image" of the demonstrator's action (if visible). The retinal image is determined according to a predefined mapping from actions to retinal binary patterns which remains fixed throughout the simulation.[3] In time steps wherein the world state or demonstrator are not visible, the corresponding input neurons are set to 0. Each of the agent's output neurons represents a motor action command, determining which actions (if any) will be executed by the agent. The output neurons (as well as the hidden neurons) are continuous neurons ranging from 0 to 1, and can thus be perceived as indicating the probability of activating each motor action. A successful agent should thus produce in each time step an activation level close or equal to 1 in the motor neuron that corresponds to the appropriate action,

[3]The selected retinal representation is of no specific significance, however, we use the representation illustrated in Figure 4.1 (wherein each action is represented by a multi-bit configuration) rather than a trivial one (wherein each action is represented by a single bit) to examine the emergence of internal localized representation of complex input patterns.

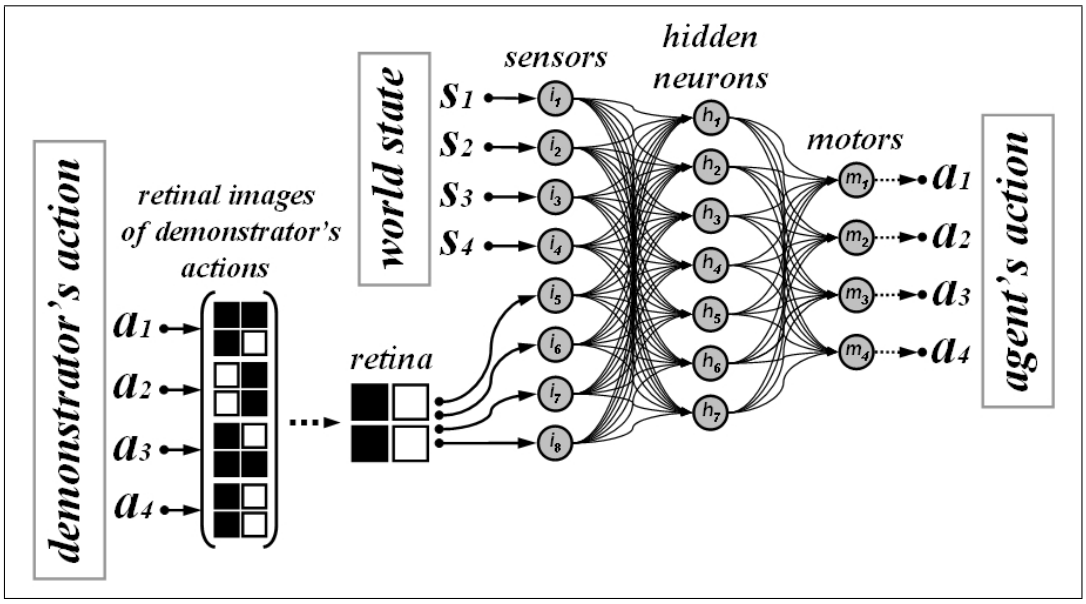


Figure 4.1: *The agent's sensorimotor system and neurocontroller.* The sensory input is binary and includes the current world state and a retinal "image" of the demonstrator's action (when visible). The retinal image for each possible demonstrator's action and a retinal input example for action a_4 are illustrated. The motor output determines which actions are executed by the agent. The network synapses are adaptive and their connection strength may change during life according to the specified learning rules.

and values close or equal to 0 in the rest of the motor neurons. In time steps where the world state is not visible (and thus, no action should be performed by the agent), a successful agent should produce activity level close or equal to 0 in all motor neurons.

Considering the agent's task and the environment it inhabits, the architecture of the agent's neurocontroller should encompass several characteristics. Clearly, it should be capable of acquiring new behaviors during the agent's life to allow imitative learning. However, to support complex dynamics which may employ both a fixed component and a learned behavior, the neurocontroller should also allow a combination of innate and acquired elements. Moreover, the precise blend of innate and acquired properties should be determined through genetic evolution. An interesting architecture that satisfies these requirements has been proposed by Floreano and Urzelai (2000), and is applied with a few modifications in the model described below.

Each agent employs a simple feed-forward neural network as a neurocontroller (i.e. the agent cannot perceive its own actions). These networks however

are *adaptive*, whereby the genotype of each individual encodes not only the initial synaptic weights but also a *Hebbian learning rule* and *learning rate* for each synapse (Floreano and Urzelai, 2000). In particular, each synapse in the network, (i, j), connecting neuron j to neuron i, is encoded by 4 genes, defining the following properties:

(i) w_{ij}^0 - the initial connection strength of the synapse (real value in the range $[0, 1]$).

(ii) s_{ij} - the connection sign (1 or -1).

(iii) η_{ij} - the learning rate (real value in the range $[0, 1]$).

(iv) Δw_{ij} - the learning rule applied to this synapse.

Each synaptic weight w_{ij} is initialized with w_{ij}^0 at the beginning of the agent's life and is updated after every time step (a sensory-motor cycle) according to:

$$w_{ij}^t = w_{ij}^{t-1} + \eta_{ij}\Delta w_{ij} \quad .$$

Δw_{ij} encodes one of five learning (modification) rules (here, o_j and o_i denote the activity of the presynaptic neuron and postsynaptic neuron respectively):

(1) No learning: $\Delta w_{ij} = 0$.

(2) Plain Hebb rule: $\Delta w_{ij} = (1 - w_{ij})o_j o_i$.

(3) Postsynaptic rule: $\Delta w_{ij} = w_{ij}(-1 + o_j)o_i + (1 - w_{ij})o_j o_i$.

(4) Presynaptic rule: $\Delta w_{ij} = w_{ij}o_j(-1 + o_i) + (1 - w_{ij})o_j o_i$.

(5) Covariance rule:

$$\Delta w_{ij} = \begin{cases} (1 - w_{ij})F(o_j, o_i) & \text{if } F(o_j, o_i) > 0 \\ (w_{ij})F(o_j, o_i) & \text{otherwise} \end{cases}$$

where $F(o_j, o_i) = \tanh(4(1 - \mid o_j - o_i \mid) - 2)$.

These rules have been selected based on neurophysiological findings (i.e. they encapsulate some of the common mechanisms of local synaptic adaptation found in biological nervous systems) and were modified to satisfy some computational constraints (e.g. in this adaptation process synapses cannot change sign and their strength is kept in the range $[0, 1]$). For a detailed description of these adaptation dynamics see Floreano and Urzelai (2000). The synaptic weights can thus adapt online, during life, using the genetically specified learning scheme. The network topology is static throughout the process and for the purpose of our simulation was set to 8-7-4 (i.e., 8 input neurons, a hidden layer with 7 neurons, and 4 output neurons), with an additional threshold unit in each layer. Such *evolutionary adaptive autonomous agents*, inspired by those presented in Todd and Miller (1991) and Floreano and Urzelai (2000), demonstrate a learning process that is supervised only indirectly, through natural selection.

4.3.3 The Evolutionary Process

A population of the agents described above evolve to successfully behave in the environment. Each agent lives in the world for 500 time steps. Fitness is evaluated according to the agent's success in performing the proper action assigned to the current world state (i.e. activating only the appropriate motor neuron), according to the state-action mapping, in each time step. An agent should perform an action only if the world state is visible and regardless of the demonstrator's visibility. We use the Mean-Square Error (MSE) measure to calculate the distance between the agent's motor output (continuous values ranging from 0 to 1) and the desired output (a value of 1 for the appropriate motor neuron and 0 for the rest), averaged over the agent's life. A MSE value of 0 thus indicates a perfectly behaving agent. The agent performance during the first 100 time steps is not evaluated (infancy phase). Fitness value is then calculated as $(1-MSE)$ and averaged over 20 trial runs in the world.

The initial population is composed of 200 individuals, each assigned a randomly selected haploid genome (i.e. each individual holds one copy of the genome), encoding the initial connection weights, learning rules and learning rates. Each new generation is created by randomly selecting agents from the previous generation and allowing them to reproduce. Agents are selected according to their fitness, using linear scaling and a roulette wheel selection scheme (Mitchell, 1998). During reproduction, 2% of the genes are mutated. Connection strength genes

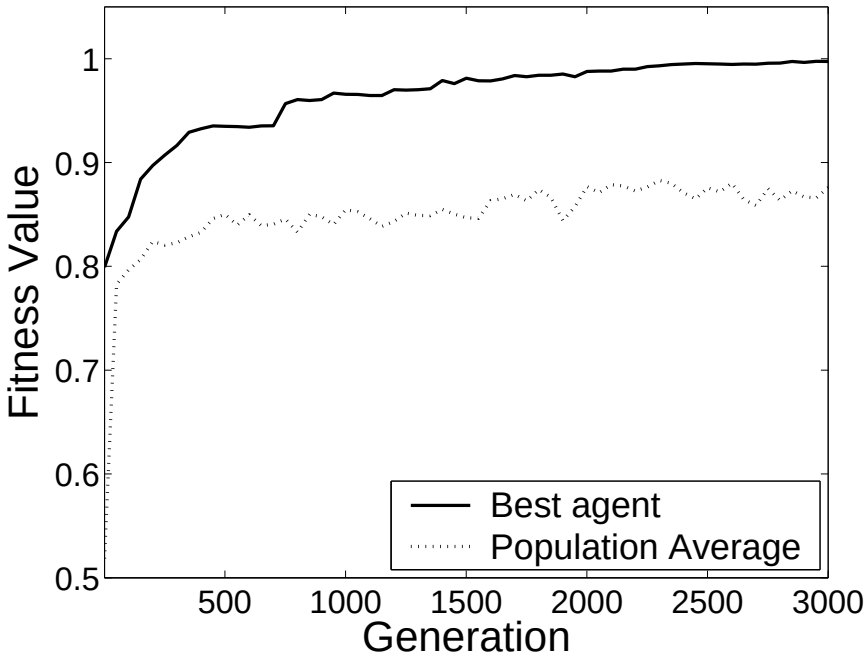


Figure 4.2: *The fitness of the best agent in the population and the population average fitness as a function of generation.*

and learning rate genes are mutated by adding a randomly selected value from the interval $[-0.3, 0.3]$, connection sign genes are mutated by flipping the sign and learning rule genes are mutated by randomly selecting one of the available rules. The genomes of the top 20% of individuals are copied to the next generation without mutation. Variations in these parameter values have no significant effect on the resulting agents.

4.4 Results and Analysis

4.4.1 The Evolution of Imitation

Within the settings described in the previous section the proper action assigned to each world state is randomly selected anew at the beginning of the agent's life. The appropriate state-action associations can thus be inferred only from the demonstrator's observed actions. Agents cannot rely on genetically coded behavior and must incorporate some sort of imitation-based learning strategy in order to demonstrate a successful behavior. Although no such learning strategy was explicitly introduced into the system, examining the fitness of the best agent in the population as a function of generation clearly demonstrates that such

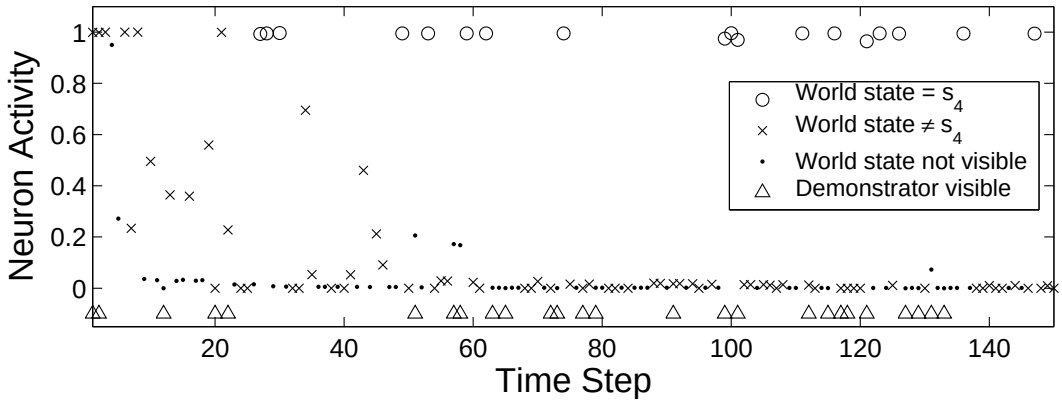


Figure 4.3: *The activation level of one motor neuron (m_2) during the first 150 time steps.* The different shapes indicate whether the world state was s_4 and whether it was visible. The triangles at the bottom further represent time steps in which the demonstrator was visible.

imitating agents have evolved (Figure 4.2). Evidently, after approximately 2000 generations, the evolved agents successfully master the behavioral task, regularly executing the proper action in each world state. Obviously, given the way the task is designed, this would not have been possible in the absence of an emerged imitation-based learning strategy.

Having successfully evolved imitating agents, we turned to examine the structure, dynamics and neural mechanisms that these agents employ. We have performed numerous evolutionary simulation runs, of which approximately half resulted in near-optimal imitating agents (exhibiting evolutionary dynamics similar to those shown in Figure 4.2). Unsuccessful simulation runs seemed to stem from early convergence of the population to sub-optimal solutions (wherein agents did not produce a distinct motor action in each time step). In the remainder of this chapter we focus on analyzing one such successful agent – the best agent in the last generation of a specific evolutionary simulation run. Other successful agents, from various evolutionary runs, were analyzed and demonstrated similar dynamics.

Direct evidence of the agent's successful imitative behavior and the resulting learning dynamics are demonstrated in Figure 4.3, depicting the activity of one of the motor neurons (m_2) in different states of the world. In this specific simulation run, the state-action mapping was arbitrarily set so that a_2 is the proper action in world state s_4 and not in any other state. In the beginning of its life, the agent activates motor m_2 (i.e., performs action a_2) whenever the world state is visible.

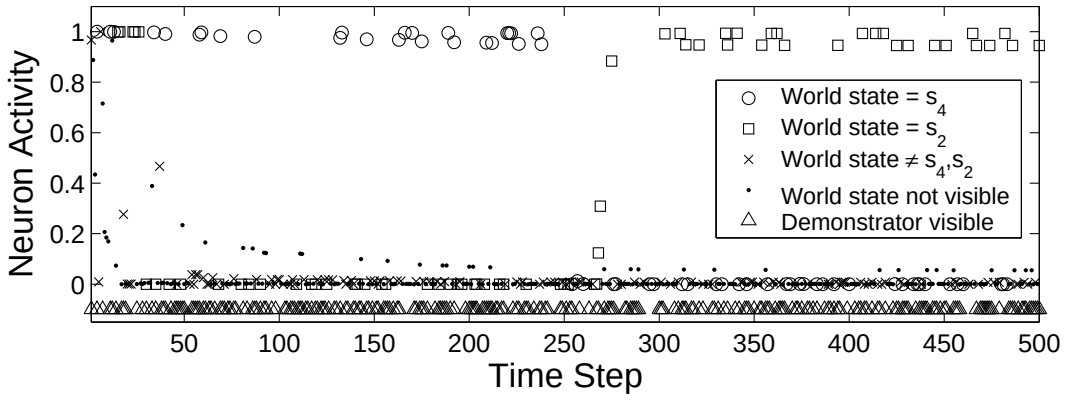


Figure 4.4: *The activation level of motor neuron m_2 during the agent's life, demonstrating the agent's ability to learn new behaviors.* In this simulation run the state-action mapping was modified in step 250, making a_2 the proper action in world state s_2 rather than s_4 as it was initially set. The triangles at the bottom further represent time steps in which the demonstrator was visible.

However, after only a few demonstrations of the appropriate behavior, the proper state-action mapping is learned and this motor is activated only when the world state is s_4, as expected. In fact, as demonstrated in Figure 4.4, the ability to learn the appropriate state-action mapping by imitation remains active during the agent's life, allowing the agent to learn a new mapping when necessary. In this experiment, the state-action mapping was initially set, as before, so that a_2 is the proper action in world state s_4. However, in the middle of the agent's life (time step 250) the state-action mapping, and accordingly the demonstrator's behavior, was changed so that a_2 is the proper action in world state s_2. Evidently, although the agent learned a certain mapping in the beginning of its life, it can quickly adapt to a new mapping after observing a few demonstrations of the new appropriate behavior.

4.4.2 The Emergence of Mirror Neurons

Examining the network hidden layer reveals an interesting phenomenon with regard to the internal representation of actions. As stated above, to support imitative learning, wherein associations from contexts to motor commands should be inferred from observations of the demonstrator's actions, an agent should be capable of matching the visual perception of an observed action to the motor command that generates the corresponding action. Figure 4.5, depicting the activation level of 3 hidden neurons, attests to the emergence of such inherent perceptual-motor

coupling. Apparently, various neurons in the hidden layer are active both when the agent performs a certain action and when it observes the demonstrator making a similar action, **forming internal mirror neurons analogous to those found in biological systems**. For the purpose of this study, we thus define mirror neurons as neurons that show a neural activation level significantly higher than 0 for both observation and execution of a certain action, and are not active in any other scenario. Although other definitions may be applied, the above definition forms a suitable analogy to the characteristics of biological neural mirroring. Interestingly, as seen in Figure 4.5, the activation level of mirror neurons during action observation is typically lower than the activation level during action execution. An analogous phenomenon can also be detected in neuronal recording data in the literature, and should be further investigated. However, in our simulation, the relatively small number of hidden neurons and mainly, the feed-forward nature of the network may account for this phenomenon, forcing mirror neurons to participate also in motor excitation.[4] These constraining properties of the artificial network, a direct consequence of several computational limitations, may also induce some constraints on the biological implications of this model, including, for example, the lack of clear distinction between active and passive perception. Such mirror neurons were found in most of the agents that evolved in our simulation environment. However, typically, not all actions in the repertoire were associated with a corresponding mirror neuron, and there have been a few cases where successful agents did not seem to incorporate any clear neural mirroring matching our above definition. There was also no evident correlation between the initial conditions or the simulation parameters and the emergence of mirror neurons, nor a clear effect of mirror neurons evolution on the fitness of the evolving agents.

The functional characteristics of the emerging mirror neurons were further examined through a set of intervention experiments, wherein hidden neurons were externally activated (stimulated) or inactivated (lesioned). These experiments confirmed that the detected mirror neurons convey the required information about the action to be performed. For example, when the world state is

[4]Furthermore, the relatively small number of hidden neurons may form a bottleneck that promotes the use of these neurons for both action perception and action execution and consequently the formation of mirror neurons. However, the fact that the same single neuron is activated in the observation and activation of the same specific action, the essence of mirroring, is surprising.

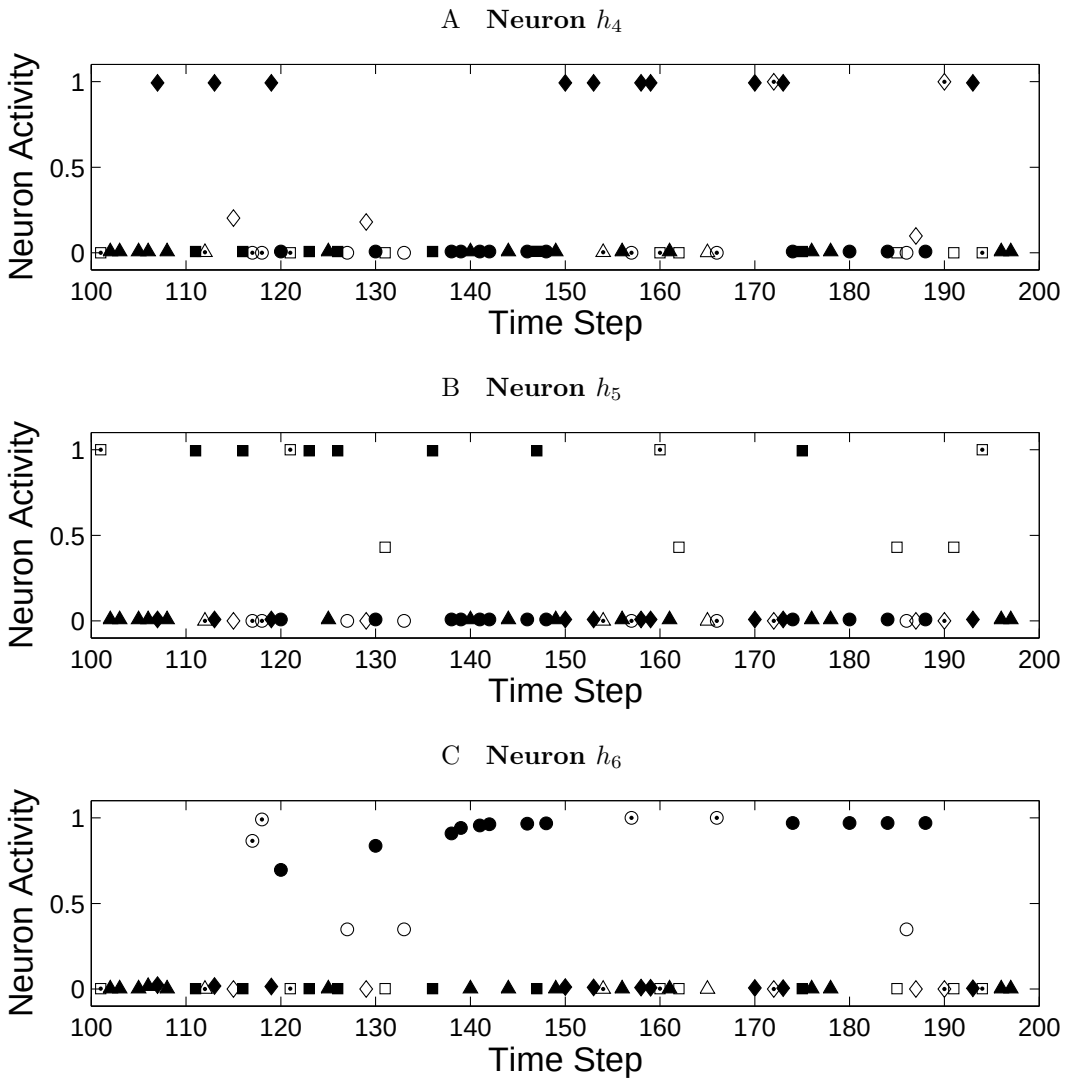


Figure 4.5: *The activation level of 3 hidden neurons (h_4, h_5 and h_6) during time steps 100-200 with an indication of the executed or observed action.* Circles, squares, diamonds and triangles represent actions a_1, a_2, a_3, a_4 respectively. An empty shape indicates that the action was only observed but not executed, a filled shape indicates that the action was executed by the agent (stimulated by a visible world state) but not observed, and a dotted shape indicates time steps in which the action was both observed and executed.

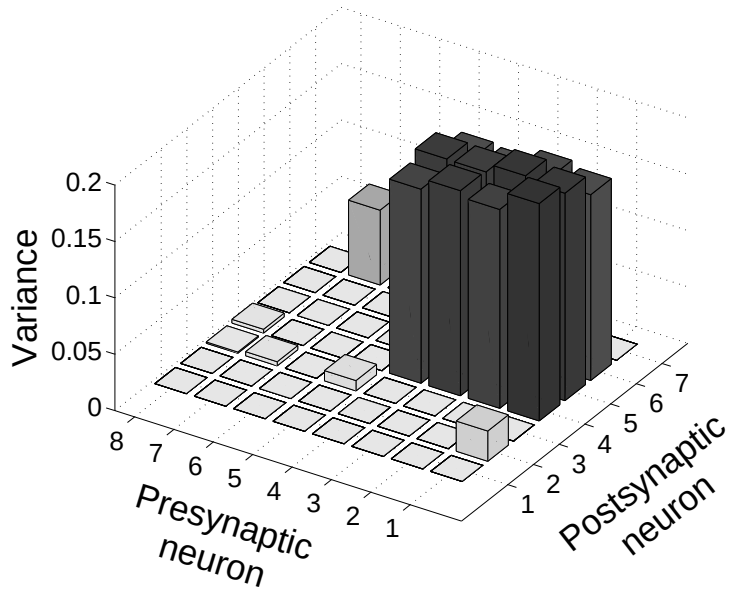
not visible (a scenario that would usually result in no action being performed) an 'artificial' stimulus of a mirror neuron resulted in the agent's performance of the action associated with that mirror neuron. Similarly, inactivating a mirror neuron inhibits the production of the associated action and in some cases resulted in the production of the wrong action.[5] Furthermore, applying *multiple* neurons activation/inactivation settings, it has been shown that even actions that could not be associated with a fully localized representation (i.e. a single mirror neuron) are still represented in the hidden layer through a distributed configuration of neurons. These findings also account for the cases mentioned above wherein successful agents did not seem to incorporate any clear localized mirror neurons. Although we cannot characterize the conditions leading to the emergence of localized mirror neurons rather than a distributed representation and although mirror neurons do not seem to have an effect on the fitness of the agent, we believe that the described setup dramatically increases the probability of mirror neuron evolution.

4.4.3 The Developmental Dynamics

We finally turn to examine the ontogenic, developmental aspects of the resulting neurocontroller. Our main objective is to identify which components in the neural mechanism are innate and which are acquired during the agent's life. We first determine which synapses play a significant role in the *learning* process. Clearly, variation in the synapse strength during life or the genetically coded learning rate are not appropriate indicators as they cannot differentiate between learning processes that genuinely adapt the agent to the world and unrelated self-organization processes. We thus measure the variance in the connection strength *at the end of the agent's life across 1000 simulation runs* (i.e. the particular agent that was analyzed above, living 1000 different lifetimes). A low variance value indicates that the synapse dynamics are independent of the world characteristics (e.g. the state-action mapping), and thus cannot contribute to the learning process that

[5]Recent reversible inactivation studies (Fogassi et al., 2001) demonstrated a distinction between two sectors in area F5 in monkeys: Mirror neurons are located in sector *F5 convexity*. Canonical neurons (neurons that respond to the presentation of three-dimensional objects of different size and shape) are located in sector *F5 bank*. While inactivation of area F5 bank produced a severe deficit of the required actions, inactivation of the cortical convexity determined only a motor slowing, preserving the appropriate action production. Clearly, within our simple model, such distinction between canonical and mirror neurons could not have developed and the mirror neurons that have emerged play a crucial role in the visuomotor pathway.

A



B

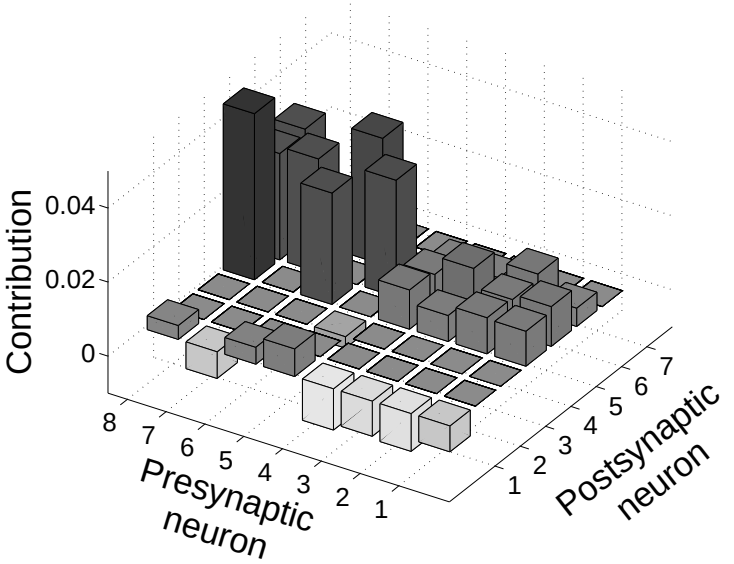


Figure 4.6: *An illustration of the connection strength variance (***A***) and the overall contribution (***B***) of the synapses connecting the sensory input layer (presynaptic) to the hidden layer (postsynaptic).* Neurons 1-4 of the presynaptic input layer represent the world state while neurons 5-8 are the retinal neurons, representing the observed demonstrator's action. Neurons 4-6 of the hidden postsynaptic layer have been identified as mirror neurons.

adapt the agent to the world. As demonstrated in Figure 4.6A, this measure highlights the acquired nature of the synapses connecting the world state neurons (input neurons 1-4), with the mirror neurons we have identified (hidden neurons 4-6). **Clearly, the acquired state-action associations are induced by these synapses**. The markedly lower variance values in other synapses from this layer and in synapses connecting hidden layer neurons to motor neurons (not illustrated here), suggest that these synapses do not play an important part in the learning process and encompass the innate properties of the network. We then turned to determining the overall contribution of each synapses to the agent's successful behavior, either learned or innate. Examining the effect of numerous multiple lesion configurations, we have utilized the Multi-perturbation Shapley value Analysis (MSA), a rigorous way to determine the importance of system elements (Keinan et al., 2004). In each configuration, a set of synapses are cancelled out by setting both their initial strength and learning rate to 0. The resulting contribution of each synapse connecting the input layer to the hidden layer is illustrated in Figure 4.6B. Evidently, the synapses that have been identified above as participating in the learning process possess a non-negligible contribution value. However, the most important synapses are among those connecting the retinal neurons (input neurons 5-8), representing the observed action, with the mirror neurons (hidden neurons 4-6). **These connections manifest the strong innate associations between the visual perception of observed actions and the internal representation of these actions, developed during the evolutionary process**.

Based on the findings described above, a simple model of the mechanism that evolved in our settings to support imitative behavior can be inferred (Figure 4.7). Notably, the required perceptual-motor coupling was not explicitly engineered into the agents, but rather emerged through evolution as an *innate* property. Furthermore, to support an effective mechanism of imitation, visually perceived actions are linked to the corresponding motor commands via fully localized internal elements, representing each action, in the form of mirror neurons. The acquired context-action stimuli can then be constructed through a simple mechanism of Hebbian learning without external supervision or reinforcement signals.

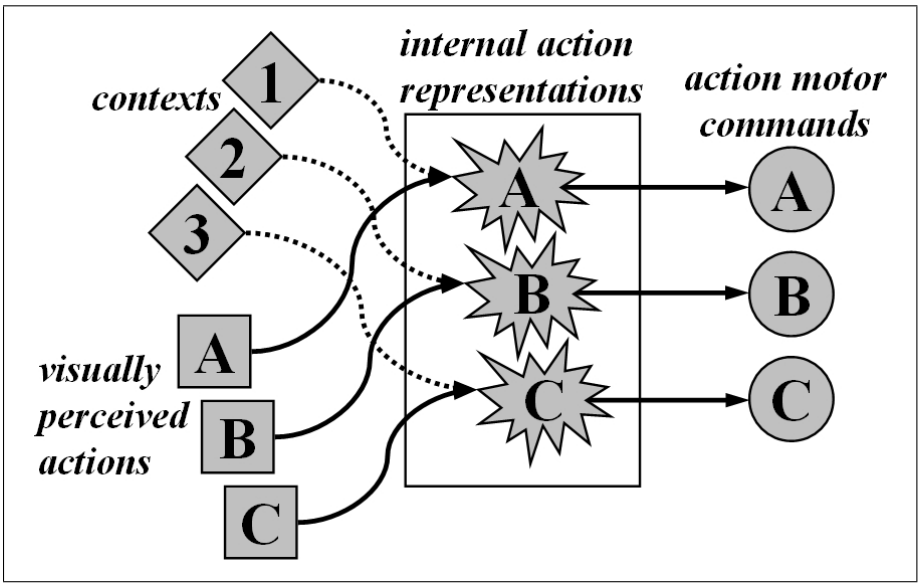


Figure 4.7: *A simple model of context-based imitation.* Solid arrows represents innate associations, while dashed arrows represents associations that are acquired during the agent's life via Hebbian learning.

4.5 Discussion

This study presents an experimental framework for studying the evolution and dynamics of imitation in evolutionary autonomous agents. This framework provides a fully accessible, distilled model for imitation and can serve as a vehicle for studying the mechanisms that underlie imitation in biological systems. As stated in Section 4.2.2, our experimental setup employs a simplified model that is not presumed to encapsulate many of the well established biological and neuronal data on imitation, nor to simulate a fully realistic social learning scenario. Rather, the aim of this model is to examine the *generic* and *universal* properties of imitative learning mechanisms. Our confidence in this framework is based on two observations: First, being an evolutionarily developed mechanism, rather than an engineered one, we believe it is likely to share the same fundamental principles driving natural systems. Second, our analysis of the resulting mechanism reveals phenomena analogous to those found in biological neural mechanisms.

The key point in our findings is that in creating a system in which only the evolution of imitation is solicited, a neural mirroring system has emerged. That is, even though no constraints on the underlying mechanisms or representations were explicitly encoded into the system, such mirror neurons have been demonstrated.

These findings imply a fundamental and essential link between the ability to imitate and a mirror system. In fact, in this regard, we believe that the simplicity of our model is one of its key assets: The emergence of neuronal mirroring to support imitation even in such a simple model, may suggest a universal and fundamental link between the ability to replicate the actions of others (imitation) and the capacity to represent and match others' actions (mirroring).

It is also important to note that although it has been hypothesized that mirror neurons underlie imitative learning functionality, the precise role of the mirror system remains unknown (Rizzolatti et al., 2001). The linkage between imitation and mirroring demonstrated in our study corroborates this hypothesis and may prove to be interesting for understanding the mechanisms that give rise to social cognitive skills. Moreover, the mirror neurons that emerged in our model, being a clear instance of a shared internal representation between observed and executed actions, also provide interesting insights that may be applied to artificial intelligence and robotic research. Although the use of internal representation is prevalent in engineered systems, the existence of such a representation in evolved systems has been challenged (Cliff and Noble, 1997). The model presented in this chapter, promoting the use of observed actions of *"others"* for learning proper motor actions of *"self"*, provides a simple example of evolved internal representation.

The framework presented in this chapter can be further enhanced to examine central issues concerning the development of imitation in animals and artifacts and the functional role of the neural mirror system. One potential application of such a framework can allow to determine the physical and social environmental conditions that promote the emergence of mirror neurons. Additionally, our framework can be enhanced to simulate a more realistic scenario of social learning. For example, one can examine how an extension of the agent's sensory input, and a complex social environment inhabited by demonstrators with varying levels of success, affect the resulting imitation strategy. Questions concerning the dependencies between observed and executed actions and the formation of neural mirroring are especially of great interest: How will the representation of actions that cannot be executed by the observer (e.g. due to different embodiment) differ from those of imitated actions? How will a hierarchical repertoire of actions affect the emerging representation? Another intriguing possibility would be to utilize this framework to explore the role of mirror neurons in the evolution of

communication (Rizzolatti and Arbib, 1998; Arbib, 2002) and in predicting the actions of others (Ramnani and Miall, 2004). It should also be noted that our experimental setup does not include agent-environment interaction and that the actions of the agent cannot affect its own next state. Focusing on the emerging representation of observed and performed actions, this choice of a non situated approach is sufficient and does not mask the obtained results with redundant factors concerning agent-environment bidirectional influence. It would be, however, interesting to also consider a situated approach, where, for example, the agent's behavior affects the actions it observes next. Specifically, one can examine whether agents in this setup could evolve a behavioral pattern that increases their future exposure and observation of novel demonstrator actions.

Furthermore, as the focus of this study is the emergence of shared and localized action representation and the matching between perceived and performed actions, our model assumes a finite and predefined set of motor actions that are already present in the action repertoire of both the imitating agent and the demonstrator. However, we do not address the question of how perceiving the action of another individual can translate into an action (of the perceiving individual) that resembles the perceived action. That is, how can an imitating individual learn a new motor action simply by observing a demonstrator? This question also relates to the "Big Five" central questions in imitation experiments presented by Dautenhahn and Nehaniv (2002a). One approach to addressing this issue applies *forward neural network models* (see, for example, Jordan and Rumelhart, 1992). These models of learning use *intentions* as inputs, which are transformed into *actions* by the learner, which, in turn, are transformed into *outcomes* by the environment. Within this approach, imitation involves learning a set of intention-outcome training pairs. Solving a learning problem is thus performed in two phases: First the learner forms a predictive internal model, transforming actions into outcomes. This model is then used to solve the mapping between intentions and actions. In the context of learning by imitation, this predictive internal model can generate predictions of the effects of executed actions and match those to the observed consequences of the demonstrator's actions, allowing an imitating individual to translate observed actions into novel actions of self.

Clearly, the simple model presented in this chapter cannot account for the full range of imitative behaviors found in nature (e.g. recognition of novel or

compound actions). However, focusing on low-level, innate imitation, this model addresses the essential questions concerning the mechanism underlying imitative behavior. It successfully demonstrates how the required associations between perceived actions, motor commands and contexts can be constructed within a hybrid adaptation process, combining evolution and lifetime learning.

Once the capacity to imitate is in place, acquired behaviors may be transmitted from one individual to the other, percolating across populations and generations and eventually facilitating cultural evolution. In the following chapter we turn to examine the resulting dynamics of such cultural evolution processes. We apply traditional population biology paradigms, adjusted and extended to encapsulate the properties of cultural evolution, for studying a model of cultural niche construction in a metapopulation and demonstrate how various social phenomena can be explained better by it.

Chapter 5

Cultural Niche Construction in a Metapopulation

Based on:

Elhanan Borenstein Jeremy Kendal and Marcus Feldman
Cultural niche construction in a metapopulation, *Theoretical Population Biology*, 70(1), 92-104, 2006.

In the previous chapter we have examined the evolutionary origins of imitative behavior, demonstrating how imitation, and specifically, the mechanisms underlying the capacity to imitate, can emerge via the evolutionary process. However, imitation is also the vehicle that eventually drives cultural evolution - the evolution of ideas, thoughts, knowledge and beliefs. Interestingly, these cultural evolution processes show many similarities to biological evolutionary dynamics and are often examined and analyzed by models originally developed within the traditional population biology discipline (e.g., Boyd and Richerson, 1985; Cavalli-Sforza and Feldman, 1981). These relatively simple mathematical models can facilitate the study of complex cultural phenomena, providing a quantitative and computational toolbox.

One such cultural evolution pattern concerns the demographic transition - a major social process occurring over the recent decades. Bongaarts and Watkins (1996) examined the correlation between the level of development[1] and the onset of the demographic transition[2] across various countries and noted a puzzling

[1]Using the human development index proposed by UNDP, 1990

[2]The demographic transition is characterized, typically, by an increase in socioeconomic development, a reduction in mortality, and a subsequent (often many years later) reduction in

phenomenon: although the earliest countries to undergo the demographic transition were those that were most highly developed, with the passage of time, the onset of the transition in different countries occurred at ever lower levels of development. They attributed this phenomenon to the influence of social interactions at varying levels (e.g., via local, national and international channels) on the diffusion dynamics of information and ideas. They presented a comprehensive interpretation of such multilevel social interactions and their effect on the adoption of fertility control, calling for further research and modeling of this phenomenon. Furthermore, evolution-minded human scientists have struggled to produce a satisfactory evolutionary explanation for the demographic transition (Borgerhoff Mulder, 1998), and to many researchers this remains a puzzle. In this respect, the analysis by Ihara and Feldman (2004) provides a potential, albeit partial, solution. In this chapter we build on this work and construct such a model in terms of a metapopulation that incorporates the concept of cultural niche construction (Odling-Smee et al., 2003; Ihara and Feldman, 2004).

The remainder of this chapter is organized as follows. We first introduce the concept of cultural niche construction and identify the need to incorporate population structure considerations into niche construction dynamics. In Section 5.2, we present the Metapopulation Cultural Niche Construction (MPCNC) model. We examine the resulting dynamics in Section 5.3, demonstrating the emergence of analogous phenomena to those described by Bongaarts and Watkins (1996), and present a local stability analysis of the equilibrium states as well as a sensitivity analysis. The chapter concludes with a discussion and a brief summary. The study presented in this chapter has been published in Borenstein et al. (2006b).

5.1 Cultural Niche Construction, Metapopulations, and Social Networks

Traditionally, biologists regarded organismic evolution as a complex dynamic process, taking place in an autonomously changing environment. However, organisms can modify their environment and thus significantly alter the selection pressures

fertility (Coale, 1974). Such a "fertility transition" is thought to result not only from rising costs and declining economic value of children (e.g., Notestein, 1953; Mace, 1996, 2000), but also from the social transmission of information concerning fertility control, and social influences, such as conformity, affecting the preference for fertility control (Cleland and Wilson, 1987; Borgerhoff Mulder, 1998).

governing the evolution of their own or other species (Lewontin, 1983). This phenomenon, termed *niche construction* (Odling-Smee et al., 2003; Laland et al., 1996, 2001), may have a profound effect on the evolutionary process, markedly changing evolutionary trajectories. Laland et al. (1996) used two-locus genetic models to investigate the evolutionary consequences of niche construction. In their models, it was assumed that a first locus governs niche-constructing behavior of individuals by affecting the amount of some resource in the environment. Viability selection, acting at a second locus, was assumed to depend on the amount of this resource. They showed that niche construction may generate selection pressures that lead to fixation of otherwise deleterious alleles, maintain genetic polymorphism where none is expected, eliminate what would otherwise be stable polymorphism, and produce time lags in the response to selection, as well as other unusual evolutionary dynamics.

For humans (and perhaps some other species), cultural traits can be analyzed in an analogous manner. In recent studies (Ihara and Feldman, 2004; Kendal et al., 2005), an analog of the two-locus model was applied to demonstrate the effect of cultural niche construction. Two culturally transmitted traits were considered where the frequency of the first vertically transmitted trait (e.g. level of education) acts as a cultural niche or background that affects the rate of oblique or horizontal transmission of a second cultural trait (e.g. adoption of fertility-reducing preferences). They showed that cultural niche construction may facilitate the '*demographic transition*' as an increase in the mean level of education facilitates, following a time lag, an increase in the preference for fertility control.

These studies assumed an unstructured population, and focused on the effects that the cultural niche may have on the cultural transmission process within that population. However, as pointed out by Bongaarts and Watkins (1996), human populations are structured, with levels of organization, hierarchy and subgroups that may markedly affect the dynamics of cultural transmission. Thus, in this study we extend the single population cultural niche construction model to incorporate population structure. We examine the process of cultural niche construction in a metapopulation (a population of populations) where the local frequency of one cultural trait in each population influences the transmission rate of a second trait both within and between the different populations. Cultural niche construction is introduced using a weighted social network that represents the level of communication within and between populations and we examine the

propagation dynamics of a second cultural trait on this network. We demonstrate how variation in the onset of cultural transitions between populations, for example the fertility transition, can be accounted for by the structure of a cultural background that facilitates social interactions between different populations.

5.2 The MPCNC Model

The Metapopulation Cultural Niche Construction (MPCNC) model considers a metapopulation of n populations and two cultural traits. The first cultural trait, E, provides a cultural background or niche that may affect the evolution of the second cultural trait, A. The frequencies of individuals in population i, and at generation t, that exhibit the combination of traits EA, Ea, eA and ea are given by $x^t_{22,i}$, $x^t_{21,i}$, $x^t_{12,i}$ and $x^t_{11,i}$, respectively, where $x^t_{22,i} + x^t_{21,i} + x^t_{12,i} + x^t_{11,i} = 1$. The frequencies of traits E and A within population i at generation t are given by $p^t_i = x^t_{21,i} + x^t_{22,i}$ and $q^t_i = x^t_{12,i} + x^t_{22,i}$, respectively. In the context of fertility transitions, the frequency of trait E might represent some index, or correlate, of human development such as the mean level of education in a population, while trait A represents the preference to adopt fertility control.

During each generation two phases of cultural transmission take place. In the first phase, traits E and A are vertically transmitted[3] (from parents to offspring) within each population i, with probabilities $b_{3,i}$, $b_{2,i}$, $b_{1,i}$ and $b_{0,i}$ for mother-father mating pairs of type E-E, E-e, e-E and e-e, respectively, and with probabilities $c_{3,i}$, $c_{2,i}$, $c_{1,i}$ and $c_{0,i}$ for mating pairs of type A-A, A-a, a-A and a-a, respectively[4] (in accordance with Cavalli-Sforza and Feldman, 1981). Note that traits e and a simply represent the default states of those that have not adopted E or A, respectively. The fertility selection coefficient, f, represents a fitness cost to parents that have adopted trait A, such that the relative number of offspring for mating pairs with traits A-A, A-a (or a-A) and a-a is $1 - f$, $1 - \frac{f}{2}$ and 1, respectively. Assuming a simple random mating scheme, the frequencies of the four types, EA, Ea, eA and ea in the offspring generation after vertical transmission (superscript,

[3]See, for example, Cavalli-Sforza et al. (1982), providing evidence for parent-offspring correlation in attitudes to education.

[4]In the case of fertility transitions, we assume that the costs and benefits associated with the adoption of a high level of education (that may affect the rate of diffusion of trait E) are subsumed within the vertical transmission coefficient, b (see Equations (5.1) and (5.2)) for each population. The validity of the assumptions underlying the dynamics of trait E is addressed in the general discussion.

v), are shown in Appendix C.1. Furthermore, assuming that for population i at generation $t = 0$, there is no statistical association between traits E and A, that is, $D_i = x_{22,i} - p_i q_i = 0$ (this assumption will be validated below), the analysis can be simplified from the recursion in the four combinations of traits (shown in Appendix C.1, equations (C.1)–(C.4)) to recursion in the two traits. Under this assumption, the frequencies of the two traits in the offspring generation after vertical transmission (indicated by superscript, t, v) are given by

$$\begin{aligned} p_i^{t,v} &= b_{3,i}(p_i^{t-1})^2 + (b_{2,i} + b_{1,i})p_i^{t-1}(1 - p_i^{t-1}) + b_{0,i}(1 - p_i^{t-1})^2 \quad (5.1) \\ W_i^{t-1} q_i^{t,v} &= (1 - f)c_{3,i}(q_i^{t-1})^2 + (1 - \frac{f}{2})(c_{2,i} + c_{1,i})q_i^{t-1}(1 - q_i^{t-1}) \\ &\quad + c_{0,i}(1 - q_i^{t-1})^2 \quad (5.2) \end{aligned}$$

The frequency of trait A in population i after fertility selection is normalized by dividing by the mean fitness in the population, $W_i^{t-1} = 1 - q_i^{t-1} f$. Throughout, we consider the simple case where $b_{3,i} = c_{3,i} = 1$, $b_{0,i} = c_{0,i} = 0$ for all populations. Also, we assume for simplicity that for each population i, there is no parental transmission bias of E, $b_{2,i} = b_{1,i} = b_i$, and that $c_{2,i} = c_{1,i} = c = 0.5$ (i.e. vertical transmission of trait A is completely unbiased in all populations). (5.1) and (5.2) then simplify to

$$\begin{aligned} p_i^{t,v} &= (p_i^{t-1})^2 + 2b_i p_i^{t-1}(1 - p_i^{t-1}) \quad (5.3) \\ q_i^{t,v} &= [(1 - f)(q_i^{t-1})^2 + (1 - \frac{f}{2})2cq_i^{t-1}(1 - q_i^{t-1})]/(1 - q_i^{t-1} f) \ . \quad (5.4) \end{aligned}$$

It should be noted that in these settings, if no other transmission phase would take place, the fitness cost of parents that have adopted trait A (i.e. $f > 0$) will prevent trait A from spreading in the population.

In the second phase, trait A is transmitted horizontally, (within the offspring generation), percolating across a social network connecting the various populations. The social network includes n vertices, each representing a population. The weight of the edge (i, j), e_{ij}, connecting vertex i to vertex j, represents the level of communication (and thus affects the transmission rate) between populations i and j. We assume that the level of communication between each pair of populations is determined by the cultural background (i.e. the frequency of trait

E) in each of these populations. In our model, the weight of each edge (i, j) at generation t is given by

$$e_{ij}^t = \mu_0 K^2 + \mu_1 \overline{p}^{t,v\,2} + \mu_2 p_i^{t,v} p_j^{t,v} \tag{5.5}$$

where K is some constant and $\overline{p}^{t,v}$ denotes the average frequency of trait E across all populations at generation t. The first term in this expression, $\mu_0 K^2$, represents a baseline communication level that is not affected by the cultural background. Hence, when $\mu_0 > 0$ but $\mu_1 = \mu_2 = 0$, the selection on A is independent of E and no niche construction occurs. The second term, $\mu_1 \overline{p}^{t,v\,2}$, represents the effect of global niche construction that results from the mean cultural background in the metapopulation, while not distinguishing between the various populations. Applying these two terms alone reduces the model to the simpler scenario of cultural niche construction in one population (the union of all populations in the metapopulation) similar to Kendal et al. (2005). Finally, the third term, $\mu_2 p_i^{t,v} p_j^{t,v}$, represents a local niche construction effect, specific to population i and its relationship with population j in the metapopulation.

By adjusting the values of μ_0, μ_1 and μ_2, we can examine the influence of each of these metapopulation niche construction '*modes*'[5]. As will be demonstrated in the next section, in order to reveal the unique characteristics of niche construction in a metapopulation (as opposed to one, uniform population), the third term of this expression, representing the local niche construction mode, is necessary.

Overall, trait A is transmitted horizontally from population j to population i with a probability that is dependent on the edge weighting, e_{ij}^t, the frequency of trait A in population j and the horizontal transmission coefficient h. We also assume that individuals in population i conform to the preference of the majority in their population, i.e. preferring trait A or the default, a, according to a conformity coefficient ψ. For $\psi > 0$, the conformity term, $1+\psi(2q_i^{t,v}-1)$, ranges from $1-\psi$ (reduced transmission) for $q_i^{t,v} = 0$ to $1+\psi$ (enhanced transmission) for $q_i^{t,v} = 1$ and does not affect transmission in a balanced population, $q_i^{t,v} = 0.5$. The frequency of trait A in population i after this horizontal transmission phase is calculated by averaging the effect of the transmission of trait A from all the populations in the metapopulation to population i and is thus given by (see also

[5]The case $\mu_0 = \mu_1 = \mu_2 = 0$ induces $e_{ij} = 0$ (no communication between the different populations), and hence, no horizontal transmission. This scenario will not be examined in this analysis as under the settings described above it always results in the extinction of trait A.

Appendix C.2)

$$q_i^t = q_i^{t,v} + \frac{1}{n}\sum_{j=1}^{n}(e_{ij}^t q_j^{t,v})h(1-q_i^{t,v})[1+\psi(2q_i^{t,v}-1)] \ . \tag{5.6}$$

Note that the total probability of horizontal transmission is the sum across all edges in the social network. The frequency of trait E in population i does not change in this phase, that is, $p_i^t = p_i^{t,v}$.

Considering these transmission phases, Appendix C.3 corroborates the validity of our assumption concerning the lack of statistical association between traits E and A (i.e. $D_i = x_{22,i} - p_i q_i = 0$).

5.3 Dynamics and Analysis

5.3.1 Dynamics

We first examine the influence of trait E, the cultural niche, on the horizontal transmission of trait A at a certain generation (the index t is thus omitted). Considering the '*effective*' frequency of trait A, that is, the mean frequency of individuals across the metapopulation from whom trait A might be acquired when weighted by local social network connections, we get

$$\frac{1}{n}\sum_{j=1}^{n}[e_{ij}q_j^v] = \mu_0 K^2 \overline{q}^v + \mu_1 \overline{p}^{v2}\overline{q}^v + \frac{\mu_2}{n}\sum_{j=1}^{n}(p_j^v q_j^v)p_i^v = \alpha + \beta p_i^v \tag{5.7}$$

where both $\alpha = (\mu_0 K^2 + \mu_1 \overline{p}^{v2})\overline{q}^v$ and $\beta = \frac{\mu_2}{n}\sum_{j=1}^{n}(p_j^v q_j^v)$ are identical for all populations at generation t. The form (5.7) is then incorporated into (5.6). Hence, the effect of the cultural niche construction on the spread of trait A in population i is positively correlated with the frequency of trait E (e.g. the *education level*) in this population. This phenomenon can be conceived as an analogue of the *rich-get-richer* phenomenon (though concerning the frequencies of trait E and A rather than actual wealth), whereby populations with a high frequency of the cultural background trait E are more affected by the level of trait A across the metapopulation, and subsequently adopt this trait faster than populations with a low frequency of trait E (a similar phenomenon was found by Bongaarts and Watkins, 1996). However, as both traits also vary over time via vertical transmission, this simple analysis is not sufficient to characterize the

overall complex behavior of this model. We therefore use a detailed simulation analysis to explore the joint dynamics of the system (5.1) through (5.6).

We examine the correlation between the frequency of trait E, and the diffusion of trait A across populations. In the following simulation runs $n = 50$, $c = 0.5$, $f = 0.1$, $h = 0.3$, $\psi = 0.1$ and $K = 1$, unless otherwise stated. However, the qualitative results are robust across a wide range of parameter values. All populations start with relatively low frequencies of E and A ($p_i^0 = q_i^0 = 0.05$). The rate of vertical transmission of trait E, (i.e. b_i) in each population is randomly selected at the beginning of the simulation from a uniform distribution on the interval $[0.5, 0.6]$, allowing for the level of cultural background (i.e., frequency of E) in the various populations to diverge over time. Clearly, populations for which b_i is higher adopt the cultural background trait, E, faster (see, for example, Figure 5.1A top panel, where the curves depicting the frequency of trait E in the different populations are basically ordered according to the value of b_i in each population). Examining the spread of trait A using different values of μ_0, μ_1 and μ_2, the effect of the different modes of the metapopulation niche construction is demonstrated clearly (Figure 5.1). While both global and local niche construction result in trait A spreading across the various populations, applying the local niche construction term also induces variation in trajectories for the spread of A (Figure 5.1B)[6]. Bongaarts and Watkins (1996) have arbitrarily measured the demographic transition onset by a fall of 10 percent in fertility from its pretransitional maximum. Here, we define trait A transition as the point in time where $q_i > 0.5$; however, similar results are obtained for other values. Examining the transition onset in the various populations, Figure 5.1B demonstrates that populations with higher frequencies of trait E (induced by higher transmission coefficient values, b_i), undergo trait A transition before populations with lower trait E frequencies. However, the onset of a transition in trait A in each population occurs at lower levels of trait E over time. Assuming that the frequency of trait E is a measure of the development level in each population and trait A represents the adoption of fertility-reduction preferences, this phenomenon is similar to the one described in Bongaarts and Watkins (1996). The robustness of this phenomena is further demonstrated in Figure 5.2, illustrating the resulting dynamics under various modified parameter values.

[6]Using no niche construction (i.e. $\mu_0 = 1$, $\mu_1 = \mu_2 = 0$) results in a qualitatively similar behavior to that demonstrated for global niche construction where all the populations go through the transition at the same time.

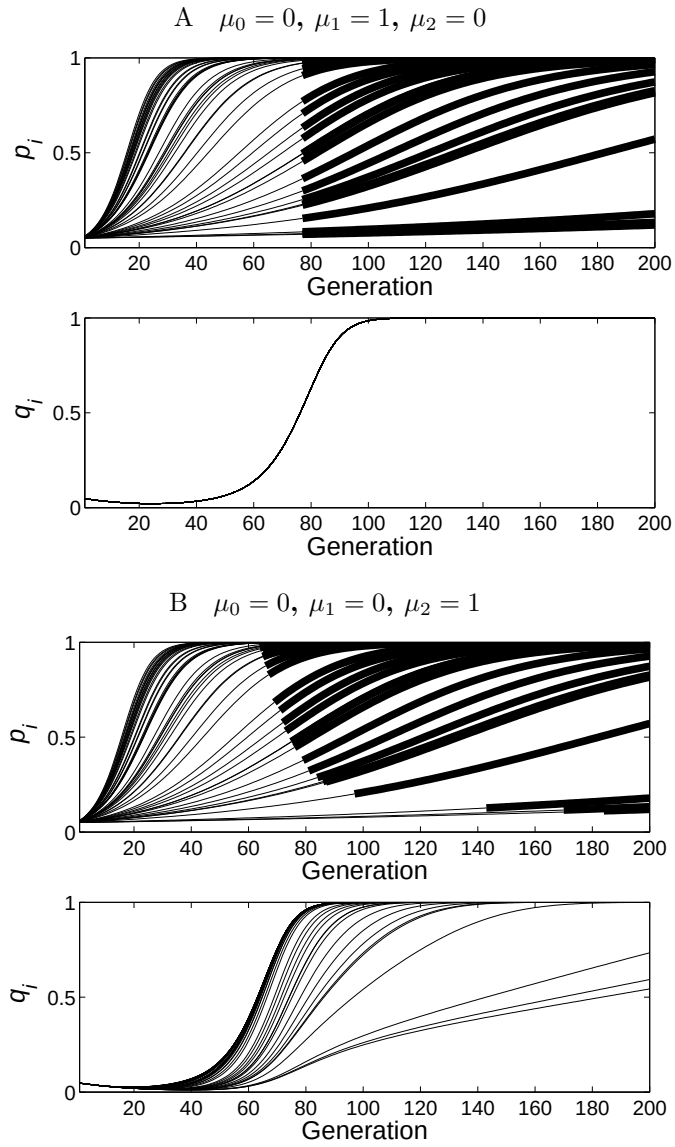


Figure 5.1: *An illustration of a simulation run of the model.* The different curves depict the frequency of traits E (top panels) and A (bottom panels) for each of the $n = 50$ populations in the metapopulation over 200 generations. Parameter values are set to $c = 0.5$, $f = 0.1$, $h = 0.3$, $\psi = 0.1$, $K = 1$ and $p_i^0 = q_i^0 = 0.05$. b_i in each population is randomly selected from a uniform distribution on the interval $[0.5, 0.6]$ (populations for which b_i is higher adopt trait E faster). In the top panel of each figure, the curves change from thin lines to thick lines when the frequency of A exceeds 0.5, indicating the onset of trait A transition (e.g. the demographic transition). (**A**) Applying global niche construction yields a uniform transition of trait A. (**B**) Local niche construction results in the various populations going through the transition at different times though at ever lower frequencies of trait E.

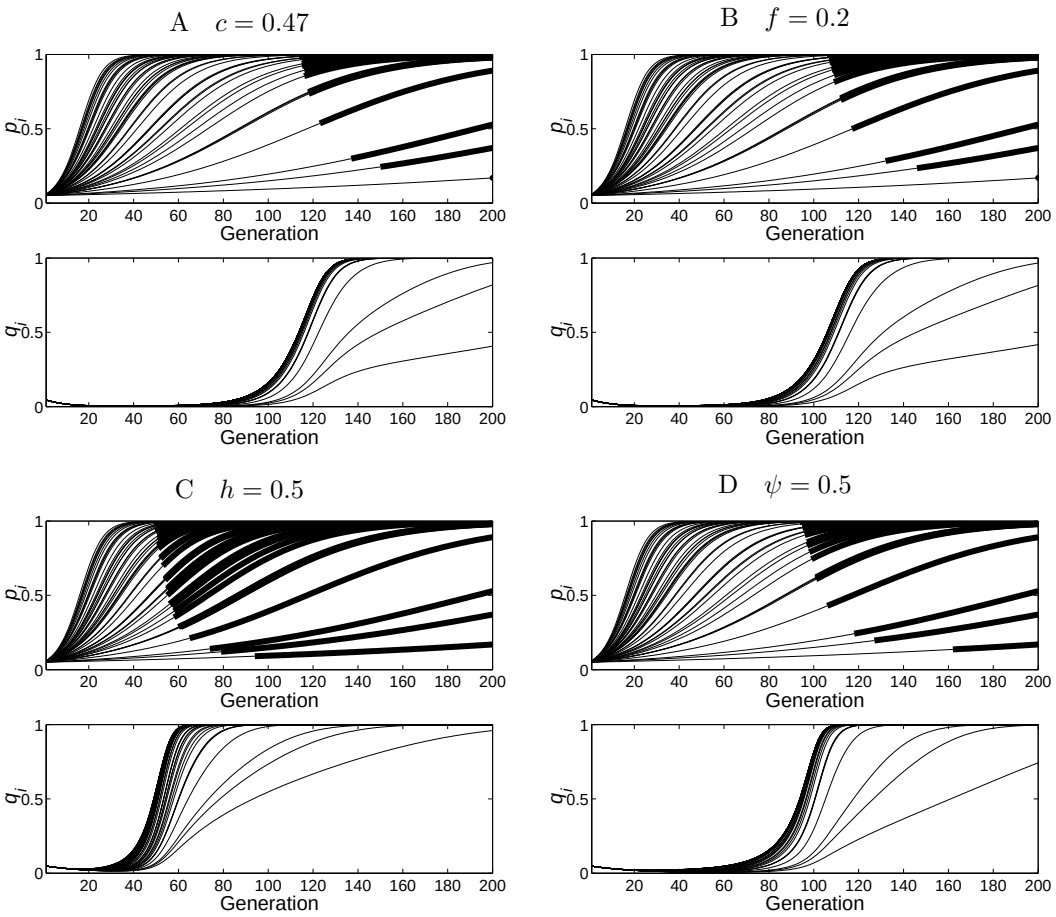


Figure 5.2: *An illustration of a simulation run of the model using various parameter values.* Aside from the modified values listed in the title of each panel, the parameters values are the same as the ones used in Figure 5.1B. Evidently, while affecting the overall spreading dynamics of trait A, all the examined parameter values still result in a qualitatively similar phenomena where the onset of a transition in trait A in each population occurs at lower levels of trait E over time.

The markedly different frequencies of trait E, at which the onset of trait A transitions occur, suggests that this effect cannot be solely attributed to varying horizontal transmission rates **within** each population. Clearly, if this were the case, different populations would have experienced the onset of transition at approximately the same level of local cultural background (i.e. $p_i \cong p_j$)[7]. To confirm the contribution of niche construction **between** populations we have conducted an additional set of simulations in which the percolation of trait A either **within** populations or **between** populations is examined in isolation. In these simulations, local niche construction is still assumed (i.e. $\mu_0 = \mu_1 = 0$, $\mu_2 = 1$), but slightly modified versions of horizontal transmission are used. These modified horizontal transmission models are given by

$$q^t_{i-within} = q^{t,v}_i + (e^t_{ii} q^{t,v}_i) h (1 - q^{t,v}_i)[1 + \psi(2q^{t,v}_i - 1)] \tag{5.8}$$

$$q^t_{i-between} = q^{t,v}_i + \frac{1}{n-1} \sum_{j=1}^{n} [(e^t_{ij} q^{t,v}_j)(1 - \delta_{ij})] h (1 - q^{t,v}_i)[1 + \psi(2q^{t,v}_i - 1)] \tag{5.9}$$

where e^t_{ij} is still calculated according to (5.5) and δ_{ij} denotes the Kronecker delta function[8]. Expressions (5.8) and (5.9) thus partition the total effect of horizontal transmission in the metapopulation (5.6) into two components, one accounting only for the horizontal transmission within each population and the other accounting only for horizontal transmission between different populations (leaving out transmission within the populations). As demonstrated in Figure 5.3, within-population niche construction indeed results in all the populations experiencing the onset of trait A transition at approximately the same frequency of trait E. Thus, the behavior demonstrated in Figure 5.1B can be attributed, almost in its entirety, to the process of local cultural niche construction between populations. In the rest of the chapter we thus revert to the original model (i.e., using Equations (5.3)–(5.6)).

The relationship between the variance in the time to reach trait A transition and the level of local niche construction effect is illustrated further in Figure 5.4. The variance is calculated for a set of simulation runs with varying degrees of local niche construction coefficient, μ_2. To maintain a similar overall communication

[7]Though not identical, as the rate of diffusion of trait E may differ across populations, resulting, over time, in variation in the influence of within-population niche construction on transition onset.

[8]The Kronecker delta is defined as having the value one when $i = j$, and zero when $i \neq j$.

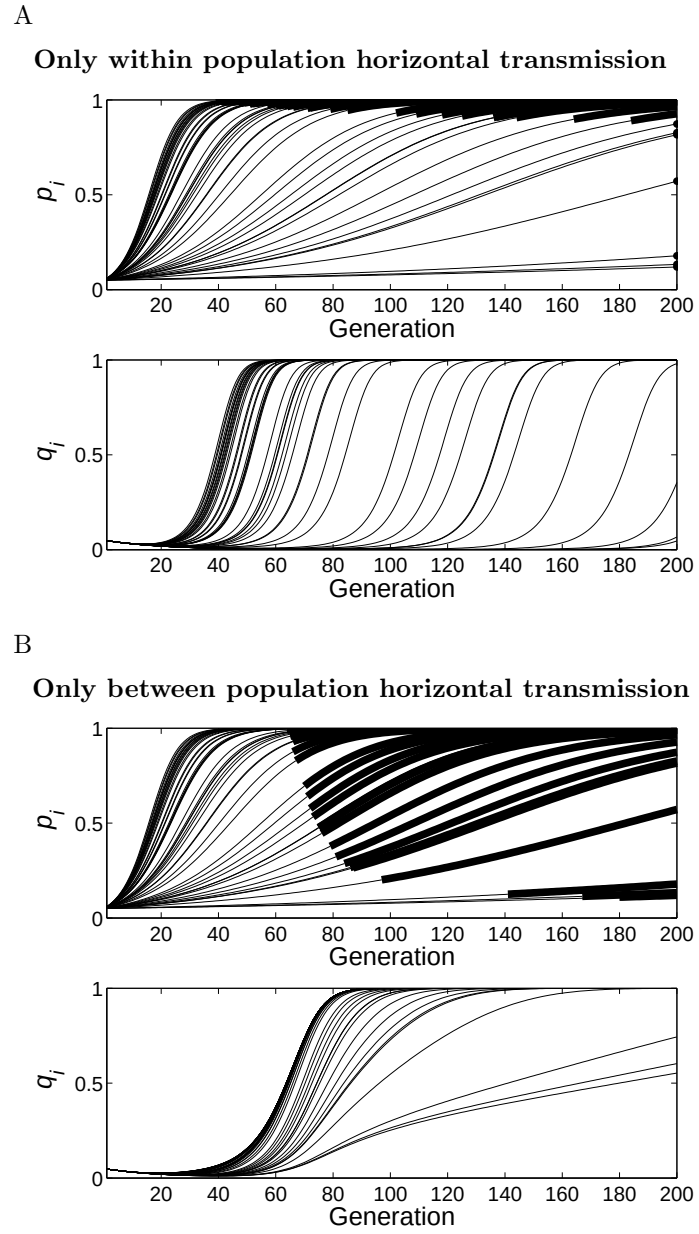


Figure 5.3: *An illustration of the effect of within population (***A***) and between population (***B***) niche construction (i.e., using Equations (5.8) and (5.9) instead of Equation (5.6)).* Aside from the modified version of horizontal transmission described above, the parameters values are the same as the ones used in Figure 5.1B.

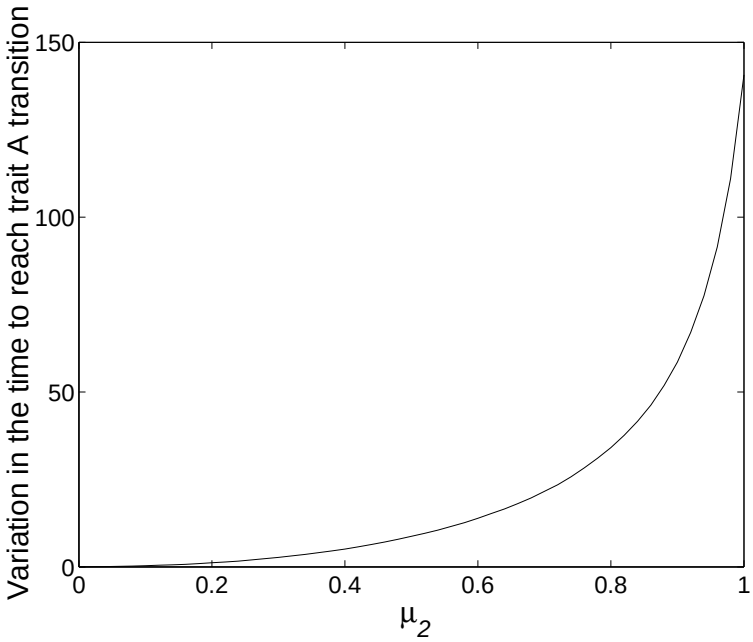


Figure 5.4: *The statistical variance of the time to reach trait A transition in the metapopulation as a function of* μ_2. The curve represents the average of 100 simulation runs. For each value of μ_2, we set $\mu_1 = 1 - \mu_2$. As demonstrated, a higher level of local niche construction yields on average higher variance.

level, μ_1 is set to $1 - \mu_2$ ($\mu_0 = 0$ in all simulations). Evidently, with higher levels of local niche construction, the variance in the onset of transitions increases. This measure could be used to estimate the relative contribution of global and local niche construction modes in human societies that exhibit this pattern of transition delay.

5.3.2 Invasion and Spread Analysis

The conditions for invasion and local stability of the four 'corner' equilibrium states, denoted by $Q_i(0,0,0)$, $Q_i(1,0,0)$, $Q_i(0,1,0)$ and $Q_i(1,1,0)$ (where $Q_i(\hat{p}_i, \hat{q}_i, D_i)$ denotes the equilibrium state), are derived from the non-collapsed version of the model (recursions (C.1)–(C.8)) and presented in Appendix C.4. In concordance with the findings of Kendal et al. (2005), it is shown that the range of parameter values under which trait A invades population i is positively related to both the coefficient of global and local niche construction. Figure 5.5 provides an example, showing the effect of global niche construction, μ_1.

We also examine the spread of trait A across the metapopulation as a function of the fertility selection, f, and the horizontal transmission coefficient, h, with

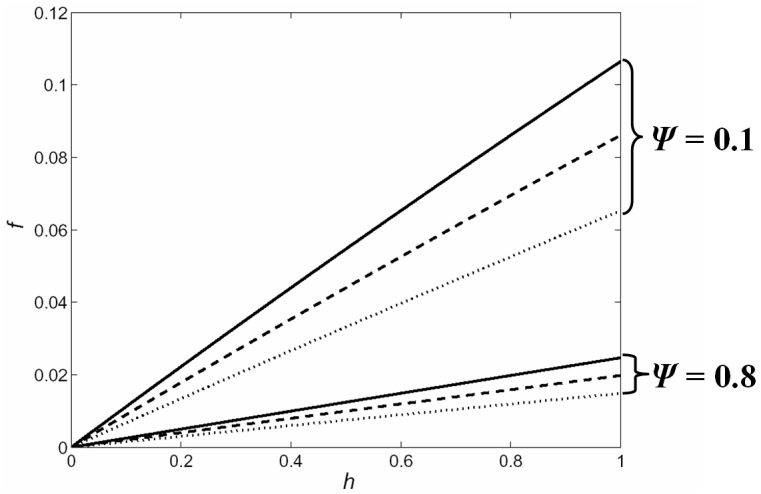


Figure 5.5: *The effect of the coefficient of conformity, ψ, and the effect of global niche construction, μ_1, on the region of invasion of trait A when perturbed from* $Q_i(0,0,0)$. $\mu_1 = 1$ (continuous line), $\mu_1 = 0.8$ (dashed line), and $\mu_1 = 0.6$ (dotted line), where, in all cases, an arbitrary value of $\overline{p} = \frac{1}{n}\sum_{i=1}^{n} p_i = 0.25$ was used. Trait A invades under each line, given by (C.19), in a single population. In particular, note that for a high level of global niche construction, trait A is predicted to invade under a lower coefficient of horizontal transmission, h, for a given level of fertility selection, f, than for a low level of global niche construction. Furthermore, global niche construction has less effect on the range of parameters for which trait A invades at high levels of conformity than at low levels of conformity. Finally, the invasion of trait A is not possible if the conformity coefficient, $\psi = 1$. Although not shown, similar trends are evident when the range of parameter values under which trait A invades population i from $Q_i(1,0,0)$ is affected by local niche construction.

varying modes of niche construction. Using a simulation run of the MPCNC model, we tested how many populations went through the trait A transition for varying levels of f and h, both from 0 to 1. The total number of generations was set to 200 to allow comparison with the dynamics described in Section 5.3.1 (other simulation parameters are set as before). As demonstrated in Figure 5.6, higher values of fertility selection require also a high level of horizontal transmission h for trait A to spread. Thus, the effects of f and h on the trait A transition are similar to their effects on the invasion of A (see, for example, Figure 5.5). A comparison between Figure 5.6A and Figures 5.6B-D shows that, typically, the conditions supporting trait A transition are less restrictive when no niche construction is applied (i.e., $\mu_1 = \mu_2 = 0$) than under either global or local niche construction. This finding, however, may be the result of the different baseline

communication level, μ_0. Note also, that in the case of $\mu_0 = \mu_1 = \mu_2 = 0$, no horizontal transmission takes place and trait A will never spread due to its fitness cost. Furthermore, as can be seen in Figure 5.6B-D, the conditions under which trait A spreads do not change significantly when the balance between global and local niche construction is varied. However, a comparison of the width of region II in Figures 5.6B-D (representing parameter values in which some populations are still in the pre-transitional state) indicates that the time to reach the transition under certain h and f parameter values increases with the coefficient of local niche construction (see also the discussion in Section 5.4).

5.3.3 Sensitivity Analysis

Here, we examine the relative sensitivity of the change in frequency of trait A in population i over horizontal transmission, $\partial\Delta q_i$, to a small change in the frequency of trait E in population j, ∂p_j, compared with a small change in the frequency of trait E within population i, given by ∂p_i. This is a measure of the relative effect of niche construction between populations i and j compared to within population i, on trait A adoption in population i. We find from (5.5) and (5.6) that

$$\frac{\frac{\partial\Delta q_i}{\partial p_j}}{\frac{\partial\Delta q_i}{\partial p_i}} = \frac{2\mu_1\overline{pq} + \mu_2 p_i q_j}{2\mu_1\overline{pq} + \mu_2(\sum_{k=1}^{n} p_k q_k + p_i q_i)}, \tag{5.10}$$

where $\overline{p} = \frac{1}{n}\sum_{j=1}^{n} p_j$, $\overline{q} = \frac{1}{n}\sum_{j=1}^{n} q_j$ and $\Delta q_i = \frac{1}{n}\sum_{j=1}^{n}(e_{ij}q_j)h(1-q_j)[1 + \psi(2q_i - 1)])$.

It is clear from (5.10) that if there is only global niche construction (i.e. $\mu_1 > 0$ and $\mu_2 = 0$),there is no difference in the sensitivity of Δq_i to a small change in the frequency of trait E within population i compared to that in population j, as all populations experience the same average network weightings (i.e. $\mu_1\overline{p}$) and thus are exposed to the same frequency of trait A (i.e. $\overline{q}$). However, if there is only local niche construction (i.e. $\mu_1 = 0$ and $\mu_2 > 0$), (5.10) is positively related to $\mu_2 p_i q_j$ (i.e., the frequency of trait A in population j that affects population i as a result of local niche construction), and is negatively related to $\mu_2\sum_{k=1}^{n} p_k q_k$ and $\mu_2 p_i q_i$ (i.e., both the total frequency of trait A across the metapopulation and the frequency of trait A within population i that affects population i as a result of local niche construction).

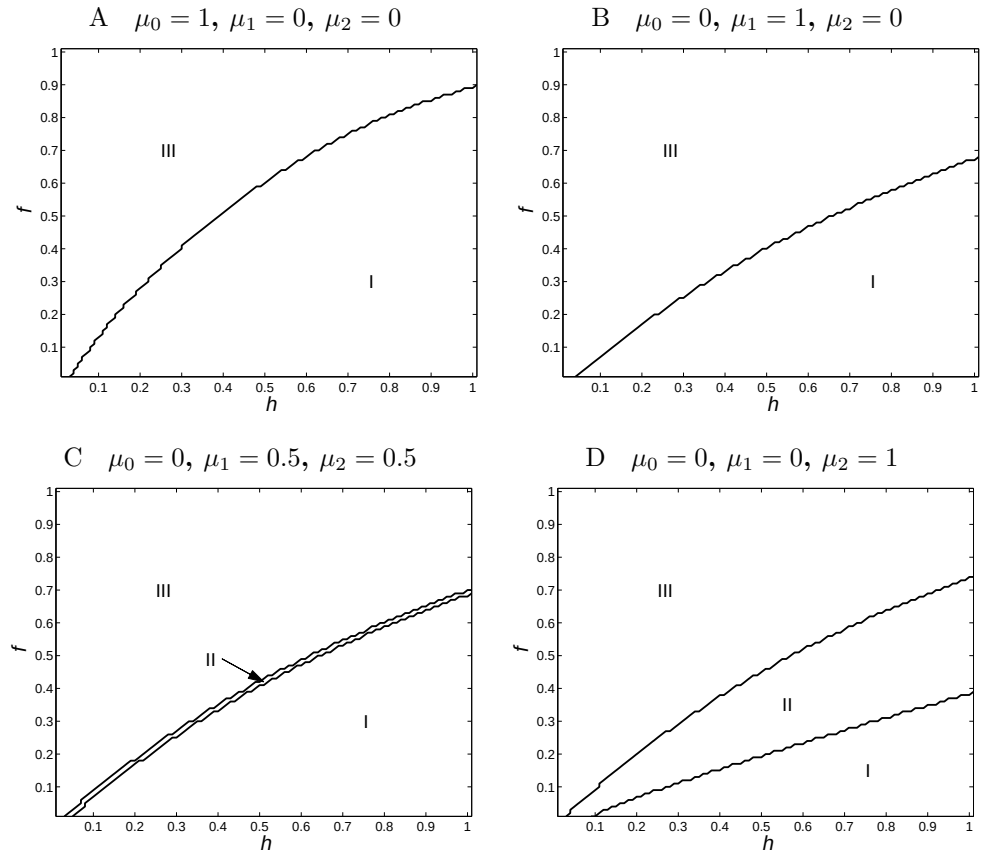


Figure 5.6: *The spread of trait A as a function of the fertility selection, f, and the horizontal transmission coefficient, h, within the first 200 generations.* Region I represents f and h values for which all 50 populations reached trait A transition (set arbitrarily as the point in time where $q_i > 0.5$), while region III indicates that the transition did not occur in *any* population. Region II represents an intermediate case where some populations are post-transitional while others are still in the pre-transitional state. Additional simulations, applying a larger number of generations, indicate that populations in region II eventually go through the transition. Comparing (**A**) with (**B-D**), it is shown that using a different baseline communication level, μ_0, modifies the conditions under which trait A spreads within 200 generations. (**B-D**) further demonstrate that varying balance between the global and local niche construction terms has relatively little effect on the spreading conditions (i.e. the regions of the h and f parameter space in which populations go to transition), although, region II in Figure 5.6D indicates that the variance in the spread of A under local niche construction may not allow all the populations to experience the transition within the first 200 generations.

If we consider the special case where the frequency of trait A is the same in all populations (i.e. for any two populations, i and j, $q_i = q_j = \overline{q}$), equation (5.10) simplifies to

$$\frac{\frac{\partial \Delta q_i}{\partial p_j}}{\frac{\partial \Delta q_i}{\partial p_i}} = \frac{2\mu_1 \overline{p} + \mu_2 p_i}{2\mu_1 \overline{p} + \mu_2 (n\overline{p} + p_i)}. \tag{5.11}$$

Figure 5.7 shows that under these conditions, sensitivity of the change in the frequency of trait A appears to be most affected by local niche construction (e.g. $\mu_1 = 0$, $\mu_2 = 1$) between populations j and i compared to within population i when the level of the cultural background in population i (or p_i) is high. This appears to be consistent with the rich-get-richer rule. Figure 5.7 further demonstrates that the same holds when the mean level of the cultural background across the metapopulation (or $\overline{p}$) is particularly low and there are few populations (n) in the metapopulation.

An additional illustration of the effect that different niche construction modes may have on the spreading dynamics of trait A is shown in Figure 5.8. In this simplified version of the model we consider only two populations ($n = 2$), and examine the change in the frequency of trait A over a single generation as a function of the frequencies of trait E in both populations using different niche construction modes. The starting point of each arrow represents the frequencies of trait E (p_1 and p_2) in the two populations. Both populations have the same arbitrary initial A frequencies, $q_1 = q_2 = 0.5$. The horizontal and vertical components of each arrow (i.e., its projections on the horizontal and vertical axes) represent the change in the frequency of trait A in both populations, Δq_1 and Δq_2 respectively, after one generation.

Clearly, in the absence of niche construction, the frequencies of trait E have no effect on the spread of trait A; $\Delta q_1 = \Delta q_2$ and they are equal for every p_1 and p_2 (Figure 5.8A). When global niche construction is applied (Figure 5.8B), the spread of trait A is influenced by the cultural background and thus Δq_1 and Δq_2 vary for different values of p_1 and p_2. For example, when p_1 and p_2 are low, the horizontal transmission rate is not sufficient to overcome the cost of fertility selection and the frequencies of trait A in both populations actually go down. Note also that the length of the arrow is the same, where $p_1 + p_2$ is constant, as all populations experience the same average network weightings and are exposed to the same frequency of trait A. Furthermore, since in global niche

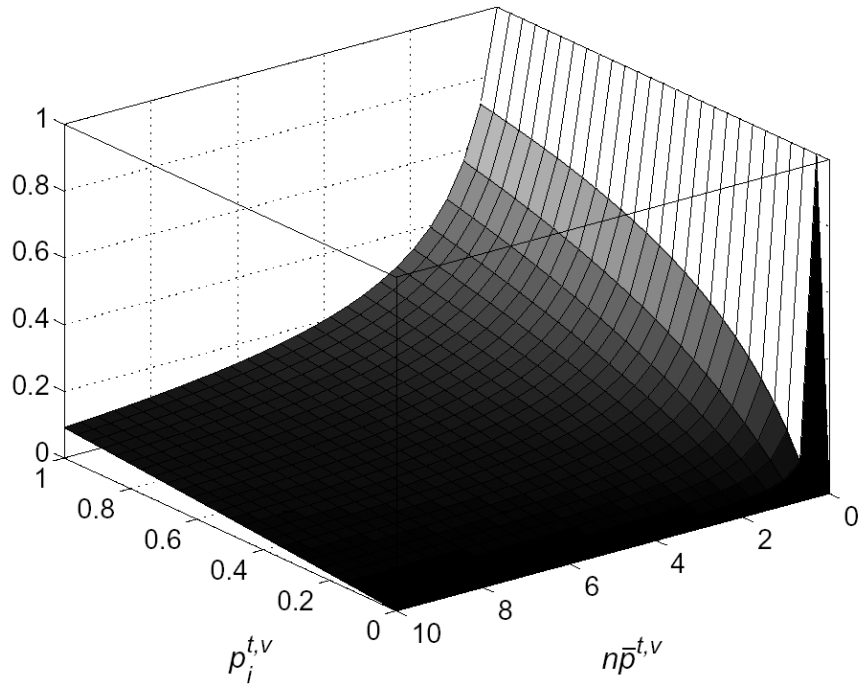


Figure 5.7: *The effect of local niche construction on the sensitivity of the change in frequency of trait A in population i to the change in frequency of trait E in population j, relative to the change in frequency of trait E within population i.* The vertical axis shows values for equation (5.11) while the other two axes show the frequency of trait E in population i, p_i, and the product of the number of populations and the mean frequency of trait E across the metapopulation, $n\overline{p}$.

construction both populations are affected similarly by the average level of the cultural background, $\Delta q_1 = \Delta q_2$, regardless of p_1 and p_2.

Only when the local niche construction mode is applied (Figure 5.8C), may the change in frequency of trait A differ between populations. Apparently, as also implied by (5.7), the spread of trait A is faster in the population with the higher level of cultural background. In fact, in some cases, we may get opposing dynamics of trait A in the two populations such that trait A spreads in one population and declines in the other (see for example the arrow for $p_1 = 0.8$, $p_2 = 0.1$). These findings again demonstrate the *rich-get-richer* phenomenon we have found in our initial analysis, whereby a higher level of trait E in one population (e.g., population 1 in the case where $p_1 > p_2$), entails a faster spread of trait A or even determines whether trait A will spread or not in that population.

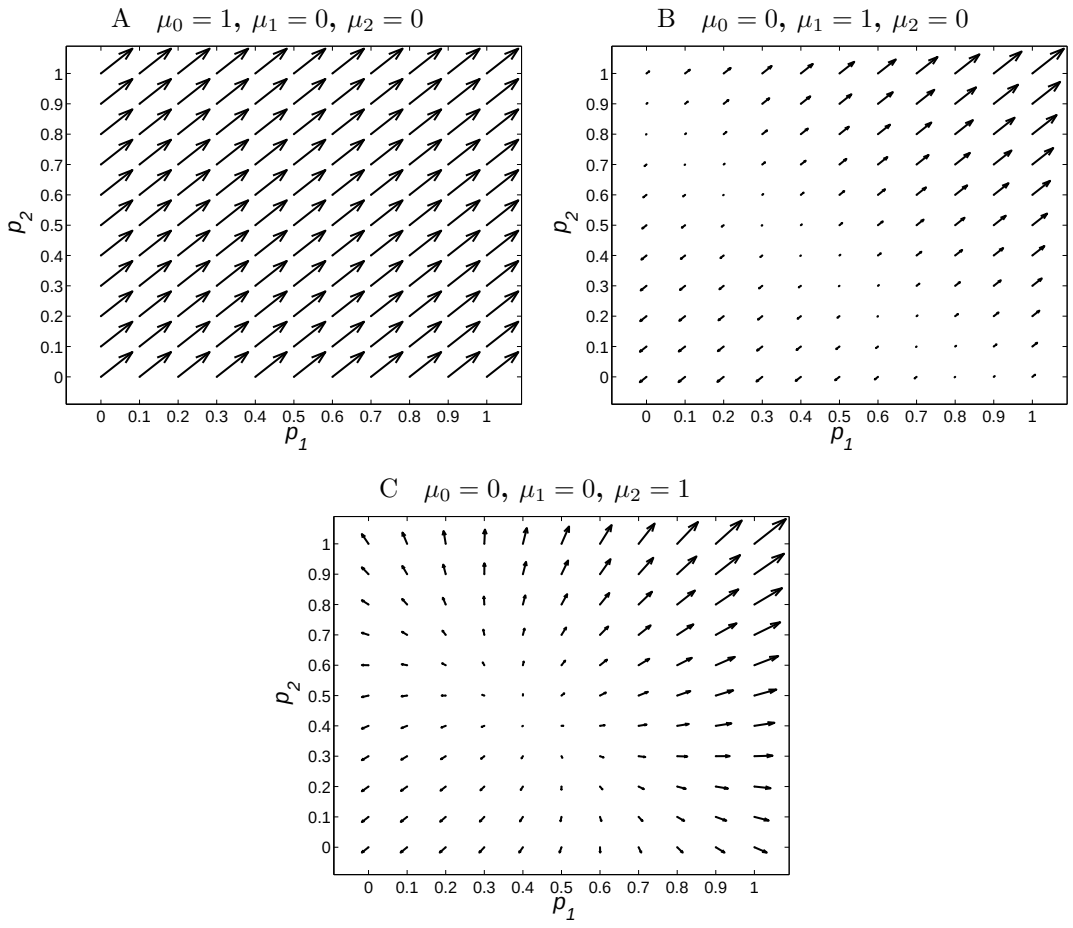


Figure 5.8: *The change in the frequency of trait A in two populations over a single generation as a function of the frequency of trait E in these populations using different modes of cultural niche construction.* The origin of each arrow represents the starting frequencies of trait E in the two populations (p_1 and p_2). In both populations, b (the vertical transmission coefficient of trait E) is set to 0.5. The horizontal and vertical components of each arrow (i.e., its projections on the horizontal and vertical axes) represent respectively the change in q_1 and q_2 in a single generation. Both populations have the same initial A frequencies, $q_1 = q_2 = 0.5$. The lengths of the arrows in each panel are scaled to fit the image grid.

5.4 Discussion

The MPCNC model examines the process of cultural niche construction in a metapopulation, extending previous models of niche construction in an unstructured population. Using a novel approach for cultural niche construction, we represent the cultural background as a dynamic social network on which other traits may percolate. This approach fuses two fundamental concepts of cultural evolution (namely, social networks and niche construction) and facilitates the analysis of markedly more realistic social interactions: Social networks, the vehicle of cultural evolution, are not static, but rather change over time, often owing to the evolutionary dynamics of other co-evolving cultural traits. As was demonstrated by our results, this manifestation of cultural niche construction should not be ignored.

Notably, analysis of our model reveals significantly different dynamics from those observed in an unstructured population facing only a global niche construction process. The spread of trait A in each population is affected not only by the level of trait E in that population but also by the level of both traits in the rest of the metapopulation. In particular, we find that introducing local niche construction results in the onset of trait A transition in each population occurring later at lower levels of trait E. The contribution of *between* (vs. *within*) population niche construction to this phenomenon has also been validated. Applying this model to the case of development (or education) and fertility control, we find that *it can account for the interesting dynamics previously reported by Bongaarts and Watkins (1996), namely, the onset of the demographic transition in different countries occurring at ever lower levels of development.*

The relative frequency of global and local niche construction does not appear to have a significant impact on the regions of the h and f parameter space in which populations go through the transition in the long term (see Figure 5.6). However, for certain h and f values, local niche construction yields a slower spread of trait A in comparison to global niche construction (see, for example, region II in Figure 5.6D), preventing some populations from going through the transition in a limited time (e.g., 200 generations). This finding coincides with our intuition that second order processes control fine-tuned details of the system whereas the overall dynamics are not affected.

Finally, our study reveals a number of interesting characteristics associated with local niche construction. For instance, local niche construction can yield

different (or even opposing) dynamics in each population, and typically appears to facilitate the rich-get-richer rule. Furthermore, the effect of local niche construction between populations, relative to within a population, is sensitive to the number of populations in the metapopulation, n.

An additional characteristic of the demographic transition is the observation that rich families reduce their fertility earlier and to a greater degree than poor families. To some extent, the analysis presented in this chapter can be interpreted as describing semi-isolated subsections of a single population, rather than multiple interacting populations. The resulting dynamics can thus also provide an explanation for this second characteristic.

The MPCNC model can be extended further to embody better the complex social interactions that exist in modern human societies. Bongaarts and Watkins (1996) describe several types of social interactions (the transmission of information and ideas, joint evaluation and social influence), but also several levels of social interaction channels (e.g., local, national and global). A more complex social network that may comprise of several hierarchies can encapsulate some of these concepts and may shed light on the importance of multi-level transmission. Similarly, there may exist more than two interacting traits, and a more elaborate interaction scheme. Furthermore, our model, as most traditional models of niche construction, assumes that the background trait evolves independently and is not affected by the niche it generates. Specifically, we assume that trait E is transferred only vertically, and does not percolate across the same social network that it itself creates. Although reasonable, it may be interesting to relax this assumption and allow a significantly more complex model where the evolution of trait E not only shapes the social network but is also affected by it (as may be the case, for example, in education over the Internet).

The model presented in this chapter further assumes that people do not migrate from one population to the other. Clearly, migration can have a strong impact on the resulting dynamics. For example, an unbiased migration (i.e. a process where people migrate randomly from one population to another) may serve as an equalizing force, reducing the variance in the frequency of the cultural traits between populations. To some extent, such dynamics may be analogous to increasing the coefficient of global niche construction (which depends only on the overall frequency of the cultural background in the metapopulation) while reducing the local niche construction coefficient. Other models may assume biased

migration where people are more likely to migrate from large populations (e.g. due to increased fertility) to smaller ones, or from less developed populations to highly developed ones. The effect of such migration schemes could be further examined through an extension of the MPCNC model that embodies this plausible process. The current model also assumes that adopting trait E does not entail any fitness cost. However, considering the case of development and fertility control, and the expenses associated with acquiring education, applying a fitness cost to the adoption of trait E may be an additional reasonable extension. In principle, however, a simple cost function should not qualitatively change the resulting dynamics induced by this model but it may slow down the spread of trait E and consequently may also hinder trait A transition.

Clearly, as extensive as the model may be, it can still encapsulate only a fraction of the real-world social interaction dynamics. However, we believe that our approach, incorporating both the cultural evolution processes and the evolution of the social network underlying these processes, can contribute to the understanding of central issues concerning the spread of cultural traits in human societies.

Chapter 6

Discussion

This dissertation presents several studies of the interaction between various lifetime adaptation mechanisms and the evolutionary process. In particular, we provide a rigorous mathematical analysis for the beneficial effect of phenotypic plasticity on evolution in multipeaked fitness landscapes and demonstrate also how a simple form of imitative learning within generations can enhance the evolution of autonomous agents. We further corroborate the fundamental evolutionary link between the capacity to imitate and the neuronal mirror system, using a novel framework of evolutionary adaptive agents. Finally, we analyze a model of cultural niche construction in a metapopulation and show how the resulting dynamics can account for complex social phenomena. The common theme of the findings presented above is the marked and bidirectional influence of the interplay between evolution and learning: Lifetime adaptation (in the form of either phenotypic plasticity or imitative learning) has been shown to dramatically accelerate the rate of the evolutionary process. Mechanisms underlying lifetime adaptation can be explained best by the evolutionary history that produced them.

This conclusion has an interesting bearing on the centuries old debate concerning *Nature vs. Nurture*, where researchers of biology, psychology, and social sciences argue about the relative importance of genetic factors vs. lifetime experience. Specifically, these researchers study how much of the individual differences in both physical and behavioral traits can be attributed to heredity and how much is determined by the environment. This debate, which can be dated back to the 13th century (Groff and McRae, 1998), remains a central topic of current research. This dissertation emphasizes that nature and nurture are strongly interlinked and that it is the interplay between them that produces many of the phenomena involved in our evolution.

In Chapter 2 we have shown that simple phenotypic plasticity schemes, including deterministic learning, stochastic learning, partial learning, and even random variation can accelerate and enhance the evolutionary process. An analogous phenomena was also observed in Chapter 3, in a scenario where such a beneficial effect is less intuitive: It was shown that imitative learning within a generation can also improve the evolutionary search although demonstrators within the population do not necessarily possess significantly improved knowledge. While the gain of lifetime learning is obvious in changing environments (allowing individuals to adapt to rapid changes that cannot be tracked by the evolutionary process), interestingly, our findings corroborate the benefit of learning and imitation also in fixed environments. We believe that such an adaptive benefit may contribute to the direct evolution of the mechanisms enabling learning and imitation. This wide-ranging advantageous effect of lifetime learning may account for the abundance of adaptation mechanisms found in nature. It would hence be interesting to examine whether a correlation between the level of plasticity displayed by various organisms and the adaptive value of learning induced by the environment they inhabit can be discovered.

There are also several fundamental principles revealed by our study. Our analysis has demonstrated the major role of the fitness landscape structure in determining the effect of plasticity on the rate of evolution. It was shown that learning indeed accelerates evolution in multipeaked landscapes but may hinder evolution in a simple unimodal scenario. Unfortunately, fully characterizing the structure of biological fitness landscapes is still far of our reach. Several promising attempts to detect structural attributes of the fitness landscape have been recently presented (Korona et al., 1994; Burch and Chao, 1999; Fong et al., 2005), providing evidence, for example, for the existence of multiple adaptive peaks. A slightly less ambitious task, which is currently at the forefront of molecular evolution research, is the characterization of the selection coefficient distribution in the vicinity of an adaptive peak (e.g., Nielsen and Yang, 2003). Some of the above studies apply a novel approach of experimental evolution of microorganisms in the lab. We hope that additional such experiments will be performed in the coming years and believe that such direct observations of an evolutionary process in progress are invaluable. Other experimental studies, focusing directly on the effect of learning on evolution, have also been recently presented (Mery and Kawecki, 2004).

Our analysis also suggests that plasticity has a beneficial effect on evolution when genotypes with low innate fitness values gain more through plasticity than genotypes with high innate fitness values, a property that is largely governed by the correlation between the genotypic and phenotypic spaces. Alas, as in the case of the fitness landscape structure, the mechanisms controlling plasticity and the adaptation dynamics they yield are still far from being completely characterized or understood.

Another interesting result stemming from the mathematical analysis presented in Chapter 2 is the increased rate of evolution in neutral regions of the landscape vs. rugged regions of positive and negative selection. It is tempting to relate this finding to the concept of neutrality (Schuster et al., 1994), robustness (de Visser et al., 2003) and designability (Kussell, 2005), studied in molecular evolution. Genetic robustness, the excess abundance of neutral mutations in a given genotype, can be conceived as the molecular evolution equivalent of the flat landscapes produced by ideal learning. It has also been shown that increased neutrality can facilitate the evolutionary process by allowing an evolving population to explore vast regions of the fitness landscape (Huynen et al., 1996b). This adaptive value of genetic robustness, induced by the structural attributes of the fitness landscape, is again analogous to the beneficial effect of lifetime learning demonstrated in our study. It has been hypothesized that the level of neutrality itself can evolve and increase during evolution (van Nimwegen et al., 1999). In a recent study (Borenstein et al., 2006a), focusing on the genetic robustness of MicroRNAs, we have provided a first biological evidence that the level of neutrality indeed increases throughout the evolutionary process and that this increase is directly selected for by evolution. Considering the analogy between genetic robustness and lifetime learning discussed above, this evidence for direct selection for robustness further strengthens our hypothesis that learning mechanisms may have been selected for by evolution owing to their beneficial effect of the landscape structure.

A major share of this dissertation addresses learning by imitation. The study of imitation has been brought back to the center of attention in recent years (Prinz and Meltzoff, 2002) and our motivation for studying imitative learning in the context of its interaction with evolution has been presented before (see Chapter 1). Our study is motivated additionally by the longstanding puzzle about how "self" and "other" are coded within our brain (Meltzoff, 1996). According to Prinz and

Meltzoff (Prinz and Meltzoff, 2002), there are two basic issues that need to be addressed by any theory of imitation: (i) How are actions perceived? (ii) How can similarity be effective between perception and action? Studying the perception and performance of actions in imitating agents from an evolutionary standpoint sheds new light on imitation in humans and primates and ultimately improves our knowledge regarding the issues raised by Prinz and Meltzoff. We believe that our research approach, focusing on the *emergence* of imitation in evolutionary autonomous agents can highlight the common underlying principles that give rise to imitative behavior. Moreover, we do not only address the questions posed above, but also tackle an additional fundamental question: "If there is indeed a qualitatively unique mechanism (or device) in the brain that can account for imitative ability, how could such a mechanism have evolved?"

Additionally, our model for the emergence of imitative learning in evolutionary adaptive autonomous agents embodies a simple, yet biologically plausible mechanism of imitative behavior, facilitating a systematic study of its structure and dynamics. The analysis of the resulting neurocontrollers reveals neural devices analogous to those found in biological systems, including clear examples of internal coupling between observed and executed actions. Further analysis of the network adaptation dynamics demonstrates the innate nature of these internal links with direct bearing on one of the key questions in imitation theory, concerning the ontogeny of mirror neurons (Meltzoff, 2002; Hurford, 2003).

Finally, the effect of cultural evolution in a metapopulation was also examined and shown to result in intriguing dynamics. Our study combines cultural evolution models, a population genetics mathematical framework and social network concepts to analyze the effect of cultural niche construction. We show how a simple model can account for puzzling social phenomena that cannot be clearly understood without considering the reciprocal effect between different populations and between several evolving cultural traits. This research also exemplifies the added-value of combining theories and models from various disciplines.

As discussed in each of the pertaining chapters, the mathematical and computational models utilized in these studies clearly cannot encapsulate the full variety and complexity of the adaptation mechanisms found in nature. Each model can be extended to incorporate additional parameters, elements, assumptions, and processes. However, focusing on the *fundamental* and *universal* properties of the interaction between learning and evolution, we believe that it is often the sim-

plicity of these models and the robustness of the results obtained that testify to their biological validity. These models form a promising test-bed for studying further central questions concerning the interplay between the various adaptation mechanisms found in nature.

Studying the dynamics of evolving populations of adaptive individuals is still far from complete. The range of mathematical and computational tools utilized in this dissertation exemplifies the challenge facing researchers in this field. We believe that the studies included in this dissertation form an important contribution to this topic. However, as in most scientific disciplines, each such study raises new and exciting research questions and only highlights how much is yet to be discovered. We hope that as more genetic, neurological and phycological data continue to accumulate, our understanding of these complex and fundamental processes, combining genetic evolution, learning, imitation and culture, will grow and evolve.

Appendix A

Mean First-Passage Times of a Simple Random Walk in an Arbitrary Environment

First, we clarify the distinction between the well established theory of Random-Walk in Random-Environments (RWRE) and our analysis: The theory of RWRE deals with the scenario where the probabilities to take a $+1$ or -1 step are independently chosen from some distribution on $(0,1)$, characterizing sub-linear or linear speed and large deviation questions. Our analysis, in contrast, relates to the simpler setup of a random walk on a finite set, but allows these probabilities to be arbitrary given constants.

A.1 One-Dimensional Random Walk: General Form and Pertaining Examples

Consider a simple random walk S_t in a changing environment on $\{0, 1, 2, \ldots, N\}$. Let $p_i = P(S_{t+1} = i+1 | S_t = i)$ and let $q_i = 1 - p_i = P(S_{t+1} = i-1 | S_t = i)$. Let $p_0 = 1$ and assume that $0 < p_i < 1$ for all $0 < i < N$. Let E_i^j denote the mean first-passage time from i to j, i.e., the expected time to first hit j starting at i. We get $E_0^1 = 1$ and the recursion

$$E_i^{i+1} = 1 + p_i \cdot 0 + (1 - p_i)[E_{i-1}^i + E_i^{i+1}] \quad , i = 1, 2, \ldots, N-1 \, . \tag{A.1}$$

Letting ρ_i denote the odds-ratio $\frac{1-p_i}{p_i}$, Eq. (A.1) may be written as

$$E_i^{i+1} = 1 + \rho_i[1 + E_{i-1}^i]$$

whose solution can be easily seen to be

$$E_i^{i+1} = 1 + 2\sum_{j=1}^{i}\prod_{k=j}^{i}\rho_k \ .$$

As $E_0^N = \sum_{i=0}^{N-1} E_i^{i+1}$ we finally obtain

$$E_0^N = N + 2\sum_{\substack{i\le j\\ 0<i,j<N}}\prod_{k=i}^{j}\rho_k \ .$$

Interestingly, this expression for the mean first-passage time from 0 to N can also be represented as the quadratic form

$$E_0^N = \mathbf{1}'\widetilde{A}\mathbf{1} \tag{A.2}$$

where

$$\widetilde{A} = \begin{pmatrix} 1 & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2\rho_3 & \rho_1\rho_2\rho_3\rho_4 & \cdots & \prod_{k=1}^{N-1}\rho_k \\ \rho_1 & 1 & \rho_2 & \rho_2\rho_3 & \rho_2\rho_3\rho_4 & \cdots & \prod_{k=2}^{N-1}\rho_k \\ \rho_1\rho_2 & \rho_2 & 1 & \rho_3 & \rho_3\rho_4 & \cdots & \prod_{k=3}^{N-1}\rho_k \\ \rho_1\rho_2\rho_3 & \rho_2\rho_3 & \rho_3 & 1 & \rho_4 & \cdots & \prod_{k=4}^{N-1}\rho_k \\ \vdots & & & & & \ddots & \vdots \\ \prod_{k=1}^{N-1}\rho_k & \prod_{k=2}^{N-1}\rho_k & \prod_{k=3}^{N-1}\rho_k & \cdots & \rho_{N-2}\rho_{N-1} & \rho_{N-1} & 1 \end{pmatrix} .$$

In particular, consider the following examples, demonstrating the mean first-passage time in a constant environment:

Example 1: Constant symmetric environment
In a constant symmetric environment, $p_i \equiv \frac{1}{2}$ ($\rho_i \equiv 1$) for all $0 < i < N$, and $\widetilde{A} = \mathbf{1}\mathbf{1}'$, the matrix of ones ($N \times N$). The quadratic form (A.2) then yields $E_0^N = N^2$.

Example 2: Constant nonsymmetric environment
In this scenario $p_i \equiv p \neq \frac{1}{2}$ ($\rho_i \equiv \rho \neq 1$) for all $0 < i < N$, and

$$\widetilde{A} = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \rho & \cdots & \rho^{N-3} \\ \vdots & & & \ddots & & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots & \rho & 1 \end{pmatrix} .$$

The quadratic form (A.2) can be easily seen to yield $E_0^N = N\frac{1+\rho}{1-\rho} + 2\rho\frac{\rho^N-1}{(\rho-1)^2}$, inducing a linear first-passage time (in N) if $\rho < 1$ and an exponential first-passage time if $\rho > 1$.

Now, consider the indices $x_1 < x_2 < \ldots < x_K$ where $\rho_{x_i} \neq 1$ ($p_{x_i} \neq \frac{1}{2}$). Define also $x_0 = 0$ and $x_{K+1} = N$. Let $n_i = x_i - x_{i-1}$ $(i = 1, 2, \ldots, K+1)$ denote the corresponding increments. As $\rho_{x_i+1}, \rho_{x_i+2}, \ldots, \rho_{x_{i+1}-1}$ all equal 1 for each $0 \leq i \leq K$, $\widetilde{A}$ consists of rectangular blocks and $E_0^N = 1'\widetilde{A}1 = V'AV$, where $V = (n_1, n_2, \ldots, n_{K+1})'$ and

$$A = \begin{pmatrix} 1 & \rho_{x_1} & \rho_{x_1}\rho_{x_2} & \rho_{x_1}\rho_{x_2}\rho_{x_3} & \rho_{x_1}\rho_{x_2}\rho_{x_3}\rho_{x_4} & \cdots & \prod_{k=1}^{K}\rho_{x_k} \\ \rho_{x_1} & 1 & \rho_{x_2} & \rho_{x_2}\rho_{x_3} & \rho_{x_2}\rho_{x_3}\rho_{x_4} & \cdots & \prod_{k=2}^{K}\rho_{x_k} \\ \rho_{x_1}\rho_{x_2} & \rho_{x_2} & 1 & \rho_{x_3} & \rho_{x_3}\rho_{x_4} & \cdots & \prod_{k=3}^{K}\rho_{x_k} \\ \rho_{x_1}\rho_{x_2}\rho_{x_3} & \rho_{x_2}\rho_{x_3} & \rho_{x_3} & 1 & \rho_{x_4} & \cdots & \prod_{k=4}^{K}\rho_{x_k} \\ \vdots & & & & & \ddots & \vdots \\ \prod_{k=1}^{K}\rho_{x_k} & \prod_{k=2}^{K}\rho_{x_k} & \prod_{k=3}^{K}\rho_{x_k} & \cdots & \rho_{x_{K-1}}\rho_{x_K} & \rho_{x_K} & 1 \end{pmatrix}.$$

Thus,

$$E_0^N = \sum_{i=1}^{K+1} n_i^2 + 2 \sum_{\substack{i \leq j \\ 0<i,j<K+1}} n_i n_{j+1} \prod_{k=i}^{j} \rho_{x_k} \,. \tag{A.3}$$

In particular, consider the following example, compactly describing the first-passage time in a typical "flattened" ideal learning landscape:

Example 3: Symmetric random walk with K equally distant disturbances

In this scenario we get $x_i = i\frac{N}{K+1}$ and $n_i \equiv n = \frac{N}{K+1}$ for all $0 < i \leq K+1$. Eq. (A.3) then yields

$$E_0^N = \frac{N^2}{K+1} + 2\frac{N^2}{(K+1)^2} \sum_{\substack{i \leq j \\ 0<i,j<K+1}} \prod_{k=i}^{j} \rho_{x_k} \,.$$

A.2 One-Dimensional Random Walk: An Asymptotic Bound

We will now turn to analyze an asymptotic bound for E_0^N. Define the "**drawdown**" R of the random walk process as the maximal element of A, that is, $R = \max_{\substack{i \le j \\ 0 \le i,j < K+1}} \prod_{k=i}^{j} \rho_{x_k}$. We want to find bounds on E_0^N in terms of R. Clearly, $\max(N, 2R) \le E_0^N \le N^2 R$. In fact, as we shall now see, $E_0^N \le N^2(\frac{1+R}{2})$ and this bound is in some sense asymptotically sharp. We first consider the case where $R > 1$ and analyze the quadratic form $V'AV$ obtained above. Fix all K values $\rho_i \neq 1$ as given, but allow $Y = (y_1, y_2, \ldots, y_{K+1}) = \frac{1}{N}(n_1, n_2, \ldots, n_{K+1})$ to vary in its feasible set $FS_N = \{(y_1, y_2, \ldots, y_{K+1}) | y_i \ge 0,\ \sum y_i = 1,\ Ny_i \text{ positive integers}\}$, a subset of $FS_\infty = \{(y_1, y_2, \ldots, y_{K+1}) | y_i \ge 0,\ \sum y_i = 1\}$. Then $\max_{Y \in FS_N} (Y'AY) \le \max_{Y \in FS_\infty} (Y'AY)$, and the two are close to each other if $K \ll N$.

Claim: $\max_{Y \in FS_\infty} (Y'AY) = \frac{1+R}{2}$

Proof:

(i) For $\rho_i < 1$, fix all y_j except y_i and y_{i+1} and fix $y_i + y_{i+1} = T$. Then $Y'AY$ is a convex quadratic function of y_i so one of y_i and y_{i+1} should be zero. This means that all $\rho_i < 1$ should "collapse" (i.e., be at distance zero) to a neighboring $\rho_j > 1$. Furthermore, all $\rho_i < 1$ left of the leftmost $\rho_j > 1$ (be that at j_0) collapse to location zero and all $\rho_i < 1$ to the right of the rightmost $\rho_j > 1$ (be that at j_1) collapse to location N.

(ii) This leaves us with a reduced matrix $\widehat{A}$ of size $(\widehat{K}+1)\times(\widehat{K}+1)$, $\widehat{K} \le K$, with $\rho_i > 1$ for all $0 < i \le \widehat{K}$. We claim that $\max(\widehat{Y}'\widehat{A}\widehat{Y})$ is achieved by collapsing all ρ's from j_0 to j_1 into one point in the exact middle. This will give

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & R \\ R & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{1+R}{2}\,.$$

Suppose that $\max_{\widehat{Y}}(\widehat{Y}'\widehat{A}\widehat{Y})$ is achieved at a vector $(y_1, y_2, \ldots, y_{\widehat{K}+1})$ with $y_i > 0$ for each $i = 1, 2, \ldots, \widehat{K}+1$. Then this vector must achieve a zero of the derivative of the Lagrangian $\widehat{Y}'\widehat{A}\widehat{Y} - \lambda(1'\widehat{Y} - 1)$, yielding $\widehat{Y} = C \cdot \widehat{A}^{-1}1$. But if

$$\widehat{A} = \begin{pmatrix} 1 & \rho_1 & \rho_1\rho_2 & & \cdots \\ \rho_1 & 1 & \rho_2 & & \cdots \\ \rho_1\rho_2 & \rho_2 & 1 & & \cdots \\ \vdots & & & \ddots & \vdots \\ & \cdots & & \rho_{\widehat{K}} & 1 \end{pmatrix}$$

then $\widehat{A}^{-1}$ is a symmetric tri-diagonal matrix with
$\widehat{A}^{-1}(1,1) = \frac{-1}{\rho_1^2-1}$, $\widehat{A}^{-1}(\widehat{K}{+}1,\widehat{K}{+}1) = \frac{-1}{\rho_K^2-1}$,
$\widehat{A}^{-1}(i,i) = -(1 + \frac{1}{\rho_{i-1}^2-1} + \frac{1}{\rho_i^2-1})$ for all $1 < i \leq \widehat{K}$, and
$\widehat{A}^{-1}(i,i{+}1) = \widehat{A}^{-1}(i{+}1,i) = \frac{\rho_i}{\rho_i^2-1}$ for all $1 \leq i \leq \widehat{K}$.

So
$(\widehat{A}^{-1}1)(1) = \widehat{A}^{-1}(1,1) + \widehat{A}^{-1}(1,2) = \frac{-1}{\rho_1^2-1} + \frac{\rho_1}{\rho_1^2-1} > 0$,
$(\widehat{A}^{-1}1)(\widehat{K}{+}1) = \widehat{A}^{-1}(\widehat{K}{+}1,\widehat{K}) + \widehat{A}^{-1}(\widehat{K}{+}1,\widehat{K}{+}1) = \frac{\rho_{\widehat{K}}}{\rho_{\widehat{K}}^2-1} + \frac{-1}{\rho_K^2-1} > 0$,
but
$(\widehat{A}^{-1}1)(i) = \widehat{A}^{-1}(i,i-1) + \widehat{A}^{-1}(i,i) + \widehat{A}^{-1}(i,i+1) = \frac{\rho_{i-1}}{\rho_{i-1}^2-1} - (1 + \frac{1}{\rho_{i-1}^2-1} + \frac{1}{\rho_i^2-1}) + \frac{\rho_i}{\rho_i^2-1} < 0$ for all $1 < i < \widehat{K}{+}1$.

As we see, $\widehat{K} = 1$ is the only feasible value and the problem reduces to

$$\begin{pmatrix} y_1 & 1-y_1 \end{pmatrix} \begin{pmatrix} 1 & R \\ R & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ 1-y_1 \end{pmatrix} = 1 + 2(R-1)y_1(1-y_1)$$

which for $R > 1$ has a maximum at $y_1 = \frac{1}{2}$, as claimed.

In the case of $R = 1$, let all $\rho_i < 1$ stick to 0 or N as before. We obtain a flat environment, for which $E_0^N = N^2$.

So, the final answer is

$$E_0^N \leq N^2(\frac{1+R}{2})$$

and this is asymptotically sharp in the sense above, where $K \ll N$. Hence, R is the critical factor determining the passage time in a given landscape.

A.3 Multidimensional Random Walk

In contrast to a one-dimensional random walk process, a multidimensional walk, propagating from a predefined starting point to another predefined end point, can take any of numerous pathways. Clearly, the characteristics of each pathway, and in particular the walk probabilities assigned to each location, determine both the dynamics of the walk along this pathway and the likelihood of choosing it. Consider all the possible pathways from this starting point to the end point on a given landscape. Each of these pathways can be conceived as a simple one-dimensional "landscape", which in turn induces a specific drawdown value. We argue that in arbitrary environments, these drawdown values will significantly differ from each other, exhibiting a roughly exponential distribution. The essential uniqueness of drawdown trajectories is similar in nature to that obtained for upper bounds by Deuschel and Zeitouni (1999) in the problem they handle, bearing some rough similarity to ours. This argument stems directly from the strong exponential effect of the pathway drift (manifested by the values of the odd-ratios ρ_i along the pathway) on the drawdown value. We will term the pathway with the minimal drawdown value 'the *Principal-Pathway*' and the drawdown value it induces 'the *Principal-Pathway drawdown*'. Considering the correlation between the first-passage time and the drawdown value in one-dimensional landscapes, we thus conjecture that all pathways apart from the Principal-Pathway are in fact irrelevant. Essentially, the random-walk process will take place along a "sausage"-like region around that minimal drawdown pathway and can be regarded as "almost" a one-dimensional process. The expected first-passage time will thus be dominated by the Principal-Pathway drawdown in an analogous manner to that shown in the one-dimensional case.

To generate rugged multidimensional fitness landscapes with a tunable drawdown value we use the following modified versions of the generalized Rastrigin and Schwefel functions, widely used multimodal benchmark functions (Mühlenbein et al., 1991; Salomon, 1996; Ballester and Carter, 2004):

$$F_{Rastrigin}(\vec{x}) = -C_r \cdot d - \sum_{i=1}^{d} x_i^2 + C_r \cdot \cos(2\pi x_i) \qquad \begin{array}{l} C_r \in [1,4] \\ x_i \in \{-5.0, -4.8, \ldots, 0\} \end{array}$$

$$F_{Schwefel}(\vec{x}) = -C_s \cdot \sum_{i=1}^{d} (-x_i \sin(\sqrt{|x_i|})) \qquad \begin{array}{l} C_s \in [0.001, 0.015] \\ x_i \in \{-500, -460, \ldots, 420\} \end{array}$$

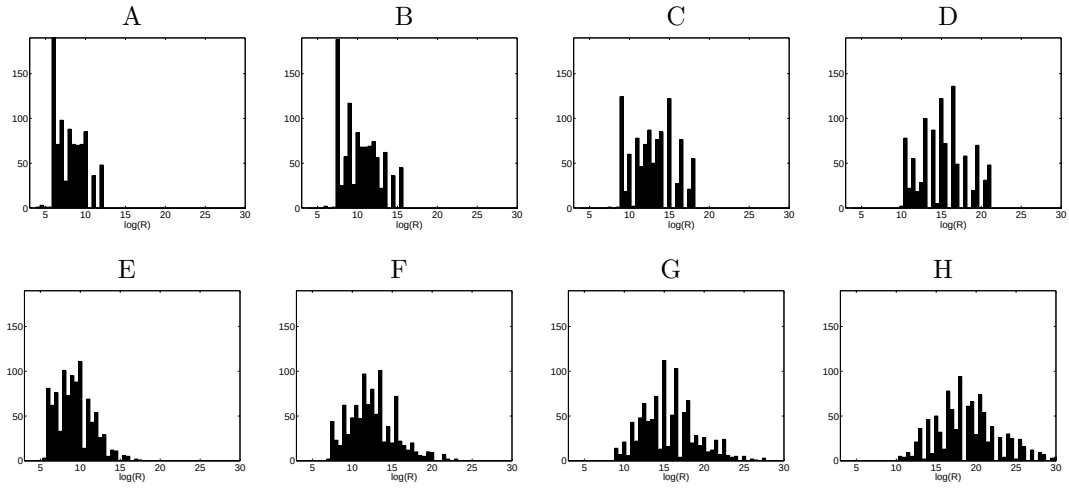


Figure A.1: *The distribution of drawdown values across randomly selected pathways on fitness landscapes with varying ruggedness.* (**A-D**) Drawdown values distribution on two-dimensional landscapes with increasing ruggedness. The landscapes are generated with the modified Rastrigin function described above with C_r coefficient of 2.4, 3.0, 3.5 and 4 respectively. The logarithm of the Principal-Pathway drawdown for these landscape is 3.702, 4.787, 5.692 and 6.596 respectively. (**E-H**) Drawdown values distribution on three-dimensional landscapes with increasing ruggedness. The C_r values and Principal-Pathway drawdowns are similar to those described in the two-dimensional case.

where d denotes the dimension (see also Figure 2.2A). These modified versions are designed so that the global optimum is located at one corner of the hypercube on which the function is defined and the starting point can be positioned on the opposite corner, at a relatively low-fitness value. C_r and C_s provide a simple way to control the ruggedness of the landscape (and consequently, the drawdown value it induces).

We validate our conjecture regarding the drawdown values distribution across various pathways numerically, using the modified Rastrigin function defined above. A set of two-dimensional and three-dimensional landscapes with varying drawdown values was generated. For each such landscape, a sample of 1000 pathways from the starting point (-5,-5) to the global optimum point (0,0) was randomly selected with uniform probability and the drawdown value of each pathway was calculated. For simplicity, we limit our sampling to forward oriented pathways (i.e. pathways composed of exactly $d(n-1)$ steps, wherein d denotes the dimension, n denotes x_i resolution in each dimension and each step is a positive step in one of the dimensions). Although actual pathways may be more complex and can

include also negative steps, this subset is sufficient to demonstrate the drawdown values distribution. Furthermore, considering only forward oriented pathways, we can calculate the Principal-Pathway drawdown using dynamic programming whereby for each point in the multidimensional space, the Principal-Pathway drawdown to that point can be obtained according to the values calculated for points that may preceded it in the pathway. Figure A.1 illustrates the resulting distribution of drawdown values for several landscapes. Evidently the variation in drawdown values induced by randomly selected pathways is extremely large (note that the distribution is drawn on a logarithm scale), and increases markedly with the ruggedness of the landscape. Moreover, as predicted above, the Principal-Pathway drawdown is markedly smaller than the drawdown of a randomly selected pathway.

Appendix B

Numerical Simulations for Studying the Effect of Phenotypic Plasticity on Evolution

B.1 The Evolutionary Process

A one-dimensional innate fitness function was defined on the interval $[1, 200]$ as a sum of several Gaussian functions, yielding a continuous, multipeaked function $F(x)$ (Figure 2.5A, solid line). Various plasticity schemes were then applied (see below) to produce the corresponding effective fitness functions.

The evolutionary process was simulated as a simple random walk. The RW probabilities in each location i are calculated as: $p_i = \mathcal{B}_T(F_i^+) = 1/[1 + e^{(F_i^- - F_i^+)/T}]$, $q_i = \mathcal{B}_T(F_i^-) = 1/[1 + e^{(F_i^+ - F_i^-)/T}]$ where $F_i^+ = F(i+1)$, $F_i^- = F(i-1)$ (using the effective fitness function in the plastic mode), and $\mathcal{B}_T(x)$ denotes the Boltzmann scaling (Mitchell, 1998) with fixed temperature $T = 0.1$. The genetic configuration x in the first generation of each evolutionary trial was set to 1.

To evaluate the convergence rate of an evolutionary process numerically, two measures are considered. First, the first-passage time of each genetic configuration provides a direct measure of the progress rate of the evolutionary process. Second, the population mean *innate* fitness value in each generation provides a good estimate for the genetic quality of the evolving individuals and allows us to track the extent of the Baldwin effect, i.e., how well did the genetically encoded solution approach the optimal one (Mayley, 1997). Clearly, evolving individuals in the plastic mode are ultimately evaluated according to their effective fitness

value (that is, their fitness after phenotypic plasticity takes place), which may be higher than their innate fitness value. However, here we focus on the rate of the evolutionary process, tracking down the genetic assimilation.

B.2 Deterministic and Stochastic Learning

Ideal deterministic learning is implemented as a simple iterative hill climbing process. Hill climbing iterations are performed until reaching the local maxima and no further improvement is possible. Partial deterministic learning was examined using a learning scheme which applies a limited number of hill climbing iterations. The resulting effective fitness functions are illustrated in Figure 2.5A.

In stochastic learning, each individual employs 100 SA (simulated annealing) iterations during its lifespan to determine its effective fitness value. Let S_i denote the individual's current configuration. In each iteration, one of the adjacent configurations (± 1) is selected at random as a candidate for a new configuration S_j. If $F(S_j) > F(S_i)$ the new configuration becomes the current configuration. Otherwise, S_j becomes the current configuration with probability $e^{(F(S_j)-F(S_i))/T}$ where T denotes a temperature parameter that cools from 1 to 0 throughout the learning process.

B.3 Random Phenotypic Variation

Given an individual with genetic configuration x, we define $F_{efc}(x) = F(x_v)$ where x_v is randomly chosen with a uniform distribution from the interval $[x - \Delta d, x + \Delta d]$. Δd thus represents the range of phenotypic variation allowed by some developmental process. This random perturbation is performed anew in each generation throughout the simulation run.

B.4 Numerical Extensions of the Random Walk Model

For the random walk with static periods extension, the process described in Section B.1 is extended to allow staying in the same genetic configuration, i, with a probability z_i. The RW probabilities in each location i are calculated as: $p_i = \mathcal{B}_T(F_i^+) = \frac{e^{F_i^+/T}}{C}$, $q_i = \mathcal{B}_T(F_i^-) = \frac{e^{F_i^-/T}}{C}$, $z_i = \mathcal{B}_T(F_i) = \frac{e^{F_i/T}}{C}$ where

$F_i = F(i)$ and C is a normalization factor (letting $p_i + q_i + z_i = 1$).

In the case of random walk with Kimura's fixation probabilities, we calculate for each genetic configuration i, the selection coefficient s for each of the two mutants ($i+1$ and $i-1$), $s_i^{\pm} = F_i^{\pm}/F_i - 1$. We then calculate the fixation probability, u, of each of the two mutants as follows: If $|s|$ is small (i.e., $|s| \leq 1/N_e$), we use Kimura's fixation probability for neutral mutations, $u = 1/(2N_e)$. Otherwise $u = (1 - e^{-4N_e sm})/(1 - e^{-4N_e s})$, where N_e denotes the effective population size and $m = 1/2N_e$ denotes the initial frequency of the mutant. In the simulations described in this study we use $N_e = 100$. We set the RW probabilities, $p_i = u^+$, $q_i = u^-$ and $z_i = 1 - p_i - q_i$.

Appendix C

MPCNC Model Recursions

C.1 Vertical Transmission

After independent vertical transmission of each of E/e and A/a, the frequency of the four cultural types in the offspring generation are given by

$$\begin{aligned} W_i x_{22,i}^{t,v} &= (1-f)c_{3,i}[b_{3,i}{x_{22,i}^{t-1}}^2 + (b_{1,i}+b_{2,i})x_{22,i}^{t-1}x_{12,i}^{t-1} + b_{0,i}{x_{12,i}^{t-1}}^2] \\ &\quad +(1-f/2)c_{2,i}[b_{3,i}x_{22,i}^{t-1}x_{21,i}^{t-1} + b_{2,i}x_{22,i}^{t-1}x_{11,i}^{t-1} + b_{1,i}x_{21,i}^{t-1}x_{12,i}^{t-1} + b_{0,i}x_{12,i}^{t-1}x_{11,i}^{t-1}] \\ &\quad +(1-f/2)c_{1,i}[b_{3,i}x_{22,i}^{t-1}x_{21,i}^{t-1} + b_{2,i}x_{21,i}^{t-1}x_{12,i}^{t-1} + b_{1,i}x_{22,i}^{t-1}x_{11,i}^{t-1} + b_{0,i}x_{12,i}^{t-1}x_{11,i}^{t-1}] \\ &\quad +c_{0,i}[b_{3,i}{x_{21,i}^{t-1}}^2 + (b_{2,i}+b_{1,i})x_{21,i}^{t-1}x_{11,i}^{t-1} + b_{0,i}{x_{11,i}^{t-1}}^2] \end{aligned} \tag{C.1}$$

$$\begin{aligned} W_i x_{21,i}^{t,v} &= (1-f)(1-c_{3,i})[b_{3,i}{x_{22,i}^{t-1}}^2 + (b_{1,i}+b_{2,i})x_{22,i}^{t-1}x_{12,i}^{t-1} + b_{0,i}{x_{12,i}^{t-1}}^2] \\ &\quad +(1-f/2)(1-c_{2,i})[b_{3,i}x_{22,i}^{t-1}x_{21,i}^{t-1} + b_{2,i}x_{22,i}^{t-1}x_{11,i}^{t-1} \\ &\qquad\qquad +b_{1,i}x_{21,i}^{t-1}x_{12,i}^{t-1} + b_{0,i}x_{12,i}^{t-1}x_{11,i}^{t-1}] \\ &\quad +(1-f/2)(1-c_{1,i})[b_{3,i}x_{22,i}^{t-1}x_{21,i}^{t-1} + b_{2,i}x_{21,i}^{t-1}x_{12,i}^{t-1} \\ &\qquad\qquad +b_{1,i}x_{22,i}^{t-1}x_{11,i}^{t-1} + b_{0,i}x_{12,i}^{t-1}x_{11,i}^{t-1}] \\ &\quad +(1-c_{0,i})[b_{3,i}{x_{21,i}^{t-1}}^2 + (b_{2,i}+b_{1,i})x_{21,i}^{t-1}x_{11,i}^{t-1} + b_{0,i}{x_{11,i}^{t-1}}^2] \end{aligned} \tag{C.2}$$

$$
\begin{aligned}
W_i x_{12,i}^{t,v} &= (1-f)c_{3,i}[(1-b_{3,i})x_{22,i}^{t-1}{}^2 + (2-b_{1,i}-b_{2,i})x_{22,i}^{t-1}x_{12,i}^{t-1} \\
&\quad +(1-b_{0,i})x_{12,i}^{t-1}{}^2] \\
&\quad +(1-f/2)c_{2,i}[(1-b_{3,i})x_{22,i}^{t-1}x_{21,i}^{t-1} + (1-b_{2,i})x_{22,i}^{t-1}x_{11,i}^{t-1} \\
&\quad +(1-b_{1,i})x_{21}^{t-1}x_{12,i}^{t-1} + (1-b_{0,i})x_{12,i}^{t-1}x_{11}^{t-1}] \\
&\quad +(1-f/2)c_{1,i}[(1-b_{3,i})x_{22,i}^{t-1}x_{21,i}^{t-1} + (1-b_{2,i})x_{21,i}^{t-1}x_{12,i}^{t-1} \\
&\quad +(1-b_{1,i})x_{22,i}^{t-1}x_{11,i}^{t-1} + (1-b_{0,i})x_{12,i}^{t-1}x_{11}^{t-1}] \\
&\quad +c_{0,i}[(1-b_{3,i})x_{21,i}^{t-1}{}^2 + (2-b_{2,i}-b_{1,i})x_{21,i}^{t-1}x_{11,i}^{t-1} \\
&\quad +(1-b_{0,i})x_{11,i}^{t-1}{}^2] \qquad \text{(C.3)}
\end{aligned}
$$

$$
\begin{aligned}
W_i x_{11,i}^{t,v} &= (1-f)(1-c_{3,i})[(1-b_{3,i})x_{22,i}^{t-1}{}^2 + (2-b_{1,i}-b_{2,i})x_{22,i}^{t-1}x_{12,i}^{t-1} \\
&\quad +(1-b_{0,i})x_{12,i}^{t-1}{}^2] \\
&\quad +(1-f/2)(1-c_{2,i})[(1-b_{3,i})x_{22,i}^{t-1}x_{21,i}^{t-1} + (1-b_{2,i})x_{22,i}^{t-1}x_{11,i}^{t-1} \\
&\quad +(1-b_{1,i})x_{21,i}^{t-1}x_{12,i}^{t-1} + (1-b_{0,i})x_{12,i}^{t-1}x_{11,i}^{t-1}] \\
&\quad +(1-f/2)(1-c_{1,i})[(1-b_{3,i})x_{22,i}^{t-1}x_{21,i}^{t-1} + (1-b_{2,i})x_{21,i}^{t-1}x_{12,i}^{t-1} \\
&\quad +(1-b_{1,i})x_{22,i}^{t-1}x_{11,i}^{t-1} + (1-b_{0,i})x_{12,i}^{t-1}x_{11,i}^{t-1}] \\
&\quad +(1-c_{0,i})[(1-b_{3,i})x_{21,i}^{t-1}{}^2 + (2-b_{2,i}-b_{1,i})x_{21,i}^{t-1}x_{11,i}^{t-1} \\
&\quad +(1-b_{0,i})x_{11,i}^{t-1}{}^2], \qquad \text{(C.4)}
\end{aligned}
$$

where $W_i = 1-f(x_{22,i}^{t-1}+x_{12,i}^{t-1}) = 1-fq_i^{t-1}$. Throughout, we consider the simple case where $b_{3,i} = c_{3,i} = 1$ and $b_{0,i} = c_{0,i} = 0$ for all populations. Also, we assume that, for population i, $b_{1,i} = b_{2,i} = b_i$ and thus any bias in vertical transmission of trait E is specific to each population, i, while assuming that $c_{1,i} = c_{2,i} = c = 0.5$, so vertical transmission of trait A is unbiased in all populations.

C.2 Horizontal Transmission

After horizontal transmission of trait A, the frequencies of the four cultural types in population i at generation $t+1$ are given by

$$
x_{22,i}^{t} = x_{22,i}^{t,v} + \frac{1}{n}\sum_{j=1}^{n} e_{ij}^{t} h(x_{12,j}^{t,v} + x_{22,j}^{t,v}) x_{21,i}^{t,v}[1 + \psi(2(x_{22,i}^{t,v} + x_{12,i}^{t,v}) - 1)] \quad \text{(C.5)}
$$

$$
x_{21,i}^{t} = x_{21,i}^{t,v} - \frac{1}{n}\sum_{j=1}^{n} e_{ij}^{t} h(x_{12,j}^{t,v} + x_{22,j}^{t,v}) x_{21,i}^{t,v}[1 + \psi(2(x_{22,i}^{t,v} + x_{12,i}^{t,v}) - 1)] \quad \text{(C.6)}
$$

$$x^t_{12,i} = x^{t,v}_{12,i} + \frac{1}{n}\sum_{j=1}^{n} e^t_{ij} h(x^{t,v}_{12,j} + x^{t,v}_{22,j}) x^{t,v}_{11,i}[1 + \psi(2(x^{t,v}_{22,i} + x^{t,v}_{12,i}) - 1)] \quad \text{(C.7)}$$

$$x^t_{11,i} = x^{t,v}_{11,i} - \frac{1}{n}\sum_{j=1}^{n} e^t_{ij} h(x^{t,v}_{12,j} + x^{t,v}_{22,j}) x^{t,v}_{11,i}[1 + \psi(2(x^{t,v}_{22,i} + x^{t,v}_{12,i}) - 1)], \quad \text{(C.8)}$$

where

$$\begin{aligned} e^t_{ij} &= \mu_0 K^2 + \mu_1 [\frac{1}{n}\sum_{j=1}^{n}(x^{t,v}_{22,j} + x^{t,v}_{21,j})]^2 + \mu_2 (x^{t,v}_{22,i} + x^{t,v}_{21,i})(x^{t,v}_{22,j} + x^{t,v}_{21,j}) \\ &= \mu_0 K^2 + \mu_1 (\overline{p}^{t,v})^2 + \mu_2 p^{t,v}_i p^{t,v}_j. \end{aligned} \quad \text{(C.9)}$$

C.3 Statistical Association between the Traits

The changes in the frequencies of traits E and A between generations $t-1$ and t in population i are given by (dropping the superscript $t-1$ on the right-hand side)

$$\Delta p_i = p_i(1-p_i)(2b_i - 1) - \frac{D_i}{W_i} f\left[b_i - p_i(2b_i - 1)\right], \quad \text{(C.10)}$$

and

$$\Delta q_i = (1-\Gamma)\frac{q_i(1-q_i)}{W_i}\left[2c(1 - f/2) - 1\right] + \Gamma(1 - q_i), \quad \text{(C.11)}$$

respectively, where $\Gamma = \frac{1}{n}\sum_{j=1}^{n}(e^t_{ij} q^{t,v}_j) h[1 + \psi(2q^{t,v}_i - 1)]$ and $D_i = x_{22,i} - p_i q_i$ denotes the statistical association between the two traits in population i.

The change in the mean frequency of traits E and A in a population, between generations t and $t+1$, is given by

$$\Delta \overline{p} = \frac{1}{n}\sum_{i=1}^{n} \Delta p_i, \quad \text{(C.12)}$$

and

$$\Delta \overline{q} = \frac{1}{n}\sum_{i=1}^{n} \Delta q_i, \quad \text{(C.13)}$$

respectively.

If, at generation $t = 0$, there is no statistical association between traits E and A in population i (i.e., $D_i^{t=0} = 0$), then there is no change in the statistical

association over time ($\Delta D_i = 0$) and D_i^t remains 0. This can be shown by noting that when $D_i^t = 0$, the frequencies of the four cultural types are given by

$$\begin{aligned} x_{22,i}^t &= p_i^t q_i^t, \\ x_{21,i}^t &= p_i^t (1 - q_i^t), \\ x_{12,i}^t &= q_i^t (1 - p_i^t) \\ x_{11,i}^t &= (1 - p_i^t)(1 - q_i^t). \end{aligned} \tag{C.14}$$

Substituting the right hand sides of (C.14) into recursions (C.1) to (C.8) reveals that $D_i^{t+1} = 0$ and thus $\Delta D_i = 0$. This makes intuitive sense as the model assumes that vertical transmission of each trait is independent of the other, while the horizontal transmission of trait A may be affected by the frequency of trait E in the population, but is not affected by whether individuals demonstrating trait A exhibit trait E or not. Also note that while $D_i = 0$, a statistical association between traits E and A over the metapopulation may result from the Wahlund effect (Wahlund, 1928). This statistical association, or covariance, is defined as $D_{metapopulation} = (\sum_i^n p_i - \overline{p})(\sum_i^n q_i - \overline{q})$ and is accounted for in the model (see equations (5.1) to (5.6)) by keeping track of p_i and q_i for each population, i.

For simulations of the model, we make the simplifying assumption that at generation $t = 0$, $D_i^{t=0} = 0$. Thus, there is never a statistical association between the two traits and so, given (C.14), recursions (C.1) to (C.8) can be collapsed to recursions for the frequency of traits E and A, over time, given by equations (5.1) to (5.6). Numerical analysis confirmed that recursions in (C.1) to (C.8) and recursions in (5.1) to (5.6) generate the same results when $D_i^{t=0} = 0$.

C.4 Local Stability of Equilibria

Here, we consider, for population i, the four 'corner' equilibrium states, denoted by $Q_i(0,0,0)$, $Q_i(1,0,0)$, $Q_i(0,1,0)$ and $Q_i(1,1,0)$ (where $Q_i(\hat{p}_i, \hat{q}_i, D_i)$ denotes the equilibrium state). From (C.10) and (C.11), the equilibrium state $Q_i(0,0,0)$ may exist if the mean frequency of trait A in the metapopulation, $\overline{q} = \frac{1}{n}\sum_{j=1}^n q_j = 0$. If $\overline{q} > 0$ and $\overline{p} = 0$, the equilibrium state $Q_i(0,0,0)$ requires that $K^2\mu_0 = 0$. Furthermore, if $\overline{q} > 0$ and, in addition, the mean frequency of trait E in the metapopulation $\overline{p} > 0$, the equilibrium state $Q_i(0,0,0)$ requires that $\mu_1 = K^2\mu_0 = 0$.

From (C.10) and (C.11), the equilibrium state $Q_i(1,0,0)$ may exist if $\overline{q} = 0$. If $\overline{q} > 0$, the equilibrium state $Q_i(1,0,0)$ requires that $K^2\mu_0 = \mu_1 = 0$. Furthermore, if $\overline{q} > 0$ and, in addition, there is a population j, for which $p_j > 0$ and $q_j > 0$, the equilibrium state $Q_i(1,0,0)$ requires that $\mu_2 = K^2\mu_0 = \mu_1 = 0$. Finally, from (C.10) and (C.11), there are two corner equilibrium states given by $Q_i(0,1,0)$ and $Q_i(1,1,0)$. When analyzing the local stability of each of the four corner equilibria in population i, we assume that the conditions required for each equilibrium to exist are satisfied and that all other populations, $j \neq i$, are at an equilibrium state.

$Q_i(0,0,0)$ and $Q_i(1,0,0)$ are unstable and invaded by trait E and by trait e if

$$b_i > 1/2, \tag{C.15}$$

and

$$b_i < 1/2, \tag{C.16}$$

respectively.

If inequalities (C.15) or (C.16) are violated, the respective equilibrium states, $Q_i(0,0,0)$ and $Q_i(1,0,0)$, are unstable and invaded by trait A when

$$2c(1-f/2) > 1/[1 + (K^2\mu_0 + \hat{\overline{p}}^2\mu_1 + \hat{p}_i^2\mu_2)h(1-\psi)]. \tag{C.17}$$

Meanwhile, the equilibrium states $Q_i(0,1,0)$ and $Q_i(1,1,0)$ are unstable and invaded by trait E and by trait e, respectively, if the respective inequalities, (C.15) and (C.16), are satisfied. If inequalities (C.15) and (C.16) are violated, the respective equilibrium states, $Q_i(0,1,0)$ and $Q_i(1,1,0)$, are unstable and invaded by trait a when

$$2(1-c)\frac{(1-f/2)}{1-f} > 1/[1 - \frac{1}{n}\sum_{j=1}^{n}(K^2\mu_0 + \hat{p}_j^2\mu_1 + \hat{p}_i\hat{p}_j\mu_2)\hat{q}_j h(1+\psi)]. \tag{C.18}$$

Inequality (C.17) can be rearranged to derive the line

$$f = \frac{2h[\mu_1\hat{\overline{p}}^2 + \mu_2\hat{p}_i^2](1-\psi)}{1 + h[\mu_1\hat{\overline{p}}^2 + \mu_2\hat{p}_i^2](1-\psi)}, \tag{C.19}$$

under which trait A may invade from either equilibrium states $Q_i(0,0,0)$ or $Q_i(1,0,0)$, where $\mu_0 = 0$ and $c = 1/2$.

Bibliography

Ackley, D. and Littman, M. (1994). Altruism in the evolution of communication. In Brooks, R. and Maes, P., editors, *Artificial Life IV: Proceedings of the International Workshop on the Synthesis and Simulation of Living Systems*. MIT Press.

Ancel, L. (1999). A quantitative model of the simpson-baldwin effect. *Journal of Theoretical Biology*, 196:197–209.

Ancel, L. (2000). Undermining the baldwin expediting effect: Does phenotypic plasticity accelerate evolution? *Theoretical Population Biology*, 58:207–319.

Ancel, L. and Fontana, W. (2000). Plasticity, evolvability, and modularity in rna. *Journal of Experimental Zoology*, 288:242–283.

Anderson, R. (1995). Learning and evolution: A quantitative genetics approach. *Journal of Theoretical Biology*, 175:89–101.

Arbib, M. (2002). The mirror system, imitation, and the evolution of language. In Dautenhahn, K. and Nehaniv, C., editors, *Imitation in Animals and Artifacts*. The MIT Press.

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review*, 80(4):1095–1111.

Baldwin, J. (1896). A new factor in evolution. *American Naturalist*, 30:441–451.

Ballester, P. and Carter, J. (2004). An effective real-parameter genetic algorithm for multimodal optimization. In Parmee, I., editor, *Adaptative Computing in Design and Manufacture VI*, pages 359–364, London. Springer.

Bargh, J. (1997). The automaticity of everyday life. In R.S. Wyer, J., editor, *The Automaticity of Everyday Life: Advances in Social Cognition*, pages 1–61. Erlbaum, Mahwah, NJ.

Belew, R. (1990). Evolution, learning, and culture: Computational metaphors for adaptive algorithms. *Complex Systems*, 4:11–49.

Belew, R. and Mitchell, M. (1996). *Adaptive Individuals in Evolving Populations: Models and Algorithms*. Addison Wesley, Reading, MA.

Best, M. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior*, 7(3/4):289–306.

Billard, A. (2000). Learning motor skills by imitation: A biologically inspired robotic model. *Cybernetics & Systems*, 32:155–193.

Billard, A. and Dautenhahn, K. (1999). Experiments in learning by imitation: Grounding and use of communication in robotic agents. *Adaptive Behavior*, 7(3/4):411–434.

Blackmore, S. (1999). *The Meme Machine*. Oxford University Press, Oxford.

Bongaarts, J. and Watkins, S. (1996). Social interactions and contemporary fertility transitions. *Population and Development Review*, 22(4):639–682.

Borenstein, E., , and Ruppin, E. (2006a). Direct evolution of genetic robustness in microRNA. *Proceedings of the National Academy of Sciences USA*, 103(17):6593–6598.

Borenstein, E., Kendal, J., and Feldman, M. (2006b). Cultural niche construction in a metapopulation. *Theoretical Population Biology*, 70(1):92–104.

Borenstein, E., Meilijson, I., and Ruppin, E. (2006c). The effect of phenotypic plasticity on evolution in multipeaked fitness landscapes. *Journal of Evolutionary Biology*, 19(5):1555–1570.

Borenstein, E. and Ruppin, E. (2003a). Enhancing autonomous agents evolution with learning by imitation. In Dautenham, K. and Nehaniv, C., editors, *Proceedings of the Second International Symposium on Imitation in Animals and Artifacts*, pages 31–38, Aberystwyth, Wales. The University of Wales.

Borenstein, E. and Ruppin, E. (2003b). Enhancing autonomous agents evolution with learning by imitation (extended version). *Journal of Artificial Intelligence and the Simulation of Behavior*, 1(4):335–347.

Borenstein, E. and Ruppin, E. (2004). Evolving imitating agents and the emergence of a neural mirror system. In Pollack, J., Bedau, M., Husbands, P., Ikegami, T., and Watson, R., editors, *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)*, pages 146–151, Cambridge, MA. MIT press.

Borenstein, E. and Ruppin, E. (2005a). The evolution of imitation and mirror neurons in adaptive agents. *Cognitive Systems Research (special issue on Epigenetic Robotics)*, 3(6):229–242.

Borenstein, E. and Ruppin, E. (2005b). The evolutionary link between mirror neurons and imitation: An evolutionary adaptive agents model. *Behavioral and Brain Sciences*, 2(28):127–128.

Borgerhoff Mulder, M. (1998). The demographic transition: Are we any closer to an evolutionary explanation? *Trends in Ecology and Evolution*, 13:266270.

Boyd, R. and Richerson, P. (1985). *Culture and the Evolutionary Process*. The University of Chicago Press, Chicago, IL.

Boyd, R. and Richerson, P. (2005). *The Origin and Evolution of Cultures*. Oxford University Press, USA.

Brodie, R. (1995). *Virus of the Mind: The New Science of the Meme*. Integral Press, Seattle, WA.

Buccino, G., Binkofski, F., Fink, G., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R., Zilles, K., Rizzolatti, G., and Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, 13:400.

Burch, C. L. and Chao, L. (1999). Evolution by small steps and rugged landscapes in the RNA virus $\phi 6$. *Genetics*, 151:921–927.

Byrne, R. and Russon, A. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21:667–684.

Cavalli-Sforza, L. and Feldman, M. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press.

Cavalli-Sforza, L., Feldman, M., Chen, K., and Dornbusch, S. (1982). Theory and observation in cultural transmission. *Science*, 4567(218):19–27.

Cleland, J. and Wilson, C. (1987). Demand theories of the fertility transition: An iconoclastic view. *Population Studies*, 41:5–30.

Cliff, D. and Noble, J. (1997). Knowledge-based vision and simple visual machines. *Philosophical Transactions of the Royal Society: Biological Sciences*, 352:1165–1175.

Coale, A. (1974). The history of the human population. *Scientific American*, 231:4051.

Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., and Martineau, J. (1998). Perception of motion and qEEG activity in human adults. *Electroencephalography And Clinical Neurophysiology*, 107:287–295.

Dautenhahn, K. (1995). Getting to know each other–artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356.

Dautenhahn, K. and Nehaniv, C. (2002a). The agent-based perspective on imitation. In Dautenhahn, K. and Nehaniv, C., editors, *Imitation in Animals and Artifacts*. The MIT Press.

Dautenhahn, K. and Nehaniv, C. (2002b). *Imitation in Animals and Artifacts*. MIT Press, Cambridge, MA.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press, New York, NY.

de Jong, G. (2005). Evolution of phenotypic plasticity: Patterns of plasticity and the emergence of ecotypes. *New Phytologist*, 166(1):101–118.

de Visser, J., Hermisson, J., Wagner, G., Ancel Meyers, L., Bagheri-Chaichian, H., Blanchard, J., Chao, L., Cheverud, J., Elena, S., Fontana, W., Gibson, G., Hansen, T., Krakauer, D., Lewontin, R., Ofria, C., Rice, S., von Dassow, G., Wagner, A., and Whitlock, M. (2003). Perspective: Evolution and detection of genetic robustness. *Evolution*, 57:1959–1972.

Demiris, Y. and Hayes, G. (2002). Imitation as a dual-route process featuring predictive and learning components: A biologically plausible computational model. In Dautenhahn, K. and Nehaniv, C., editors, *Imitation in Animals and Artifacts*. The MIT Press.

Demiris, Y. and Johnson, M. (2003). Distributed, predictive perception of actions: A biologically inspired robotics architecture for imitation and learning. *Connection Science Journal*, 15(4):231–243.

Deuschel, J. and Zeitouni, O. (1999). On increasing subsequences of I.I.D. samples. *Combinatorics, Probability and Computing*, 8(3):247–263.

Dewitt, T. and Scheiner, S. (2004). *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford University Press, New York, NY.

Dopazo, H., Gordon, M., Perazzo, R., and Risau-Gusman, S. (2001). A model for the interaction of learning and evolution. *Bulletin of Mathematical Biology*, 63:117–134.

Downing, K. (2004). Development and the baldwin effect. *Artificial Life*, 10(1):39–63.

Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, 73:2608–2611.

Fahlman, S. E. (1989). Faster-learning variations on back-propagation: An empirical study. In *Proceedings of the 1988 Connectionist Models Summer School*, Los Altos, CA. Morgan-Kaufmann.

Feldman, M. and Laland, K. (1996). Gene-culture coevolutionary theory. *Trends in Ecology and Evolution*, 11(11):453–457.

Fisher, R. (1930). *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Flentge, F., Polani, D., and Uthmann, T. (2001). Modelling the emergence of possession norms using memes. *Journal of Artificial Societies and Social Simulation*, 4(4).

Floreano, D. and Mondada, F. (1996). Evolution of plastic neurocontrollers for situated agents. In Maes, P., Mataric, M., Mayer, J., Pollack, J., and Wilson, S., editors, *From Animals to Animates*, volume IV, Cambridge, MA. MIT Press.

Floreano, D. and Urzelai, J. (1998). Evolution and learning in autonomous robotic agents. In Mange, D. and Tomassini, M., editors, *Bio-inspired Computing Systems*, pages 1–36. PPUR, Lausanne.

Floreano, D. and Urzelai, J. (2000). Evolutionary robots with on-line self-organization and behavioral fitness. *Neural Networks*, 13:431–443.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., and Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. *Brain*, 124:571–586.

Fong, S., Joyce, A., and Palsson, B. (2005). Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Research*, 15:1365–1372.

Fontanari, J. and Meir, R. (1990). The effect of learning on the evolution of asexual populations. *Complex Systems*, 4:401–414.

French, R. and Messinger, A. (1994). Genes, phenes and the baldwin effect: Learning and evolution in a simulated population. In Brooks, R. and Maes, P., editors, *Artificial Life IV*, pages 277–282, Cambridge, MA. MIT Press.

Futuyma, D. (2005). *Evolution*. Sinauer Associates, Sunderland, MA.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119:593–609.

Gavrilets, S. and Hastings, A. (1994). A quantitative-genetic model for selection on developmental noise. *Evolution*, 48:1478–1486.

Gillespie, J. (2004). *Population Genetics: A Concise Guide*. The Johns Hopkins University Press, Baltimore, MD, 2nd edition.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company, Reading, MA.

Groff, P. and McRae, L. (1998). The nature-nurture debate in thirteenth-century france. In *In Proceedings The Annual Meeting of the American Psychological Association*, Chicago, IL.

Gruau, F. and Whitley, D. (1993). Adding learning to the cellular development of neural networks: Evolution and the baldwin effect. *Journal of Evolutionary Computation*, 1(3):213–233.

Haldane, J. (1932). *The Causes of Evolution*. Longmans, Green and Co., London.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences USA*, 95:15061–15065.

Hartl, D. (2000). *A Primer of Population Genetics*. Sinauer Associates, Sunderland, MA, 3rd edition.

Hayes, G. and Demiris, J. (1994). A robot controller using learning by imitation. In *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems*.

Herman, L. (2002). Vocal, social, and self imitation by bottlenosed dolphins. In Dautenhahn, K. and Nehaniv, C., editors, *Imitation in Animals and Artifacts*. The MIT Press.

Hertz, J., Krogh, A., and Palmer, R. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley, Reading, MA.

Heyes, C. and Galef, B. (1996). *Social Learning in Animals: The Roots of Culture*. Academic Press.

Hinton, G. and Nowlan, S. (1987). How learning can guide evolution. *Complex Systems*, 1:495–502.

Hughes, B. (1995). *Random Walks and Random Environments*, volume 1. Clarendon Press, Oxford.

Hurford, J. (2003). Language beyond our grasp: What mirror neurons can, and cannot, do for language evolution. In Kimbrough Oller, U. G. and Plunkett,

K., editors, *Evolution of Communication Systems: A Comparative Approach*. MIT Press, Cambridge, MA.

Huynen, M., Stadler, P., and Fontana, W. (1996a). Smoothness within ruggedness: The role of neutrality in adaptation. *Proceedings of the National Academy of Sciences USA*, 93:397–401.

Huynen, M., Stadler, P., and Fontana, W. (1996b). Smoothness within ruggedness: The role of neutrality in adaptation. *Proceedings of the National Academy of Sciences USA*, 93:397–401.

Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286:2526–2528.

Ihara, Y. and Feldman, M. (2004). Cultural niche construction and the evolution of small family size. *Theoretical Population Biology*, 65:105–111.

Jordan, M. and Rumelhart, D. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, 16(3):307–354.

Kallel, L., Naudts, B., and Reeves, C. (2001). Properties of fitness functions and search landscapes. In Kallel, L., Naudts, B., and Rogers, A., editors, *Theoretical Aspects of Evolutionary Computing*, pages 175–206. Springer, Berlin.

Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Theoretical Biology*, 128:11–45.

Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, NY.

Kawamura, S. (1963). The process of sub-culture propagation among japanese macaques. In Southwick, C. and Nostrand, V., editors, *Primates Social Behaviour*, pages 82–90, New York, NY.

Keijzer, F. (2002). Representation in dynamical and embodied cognition. *Cognitive Systems Research*, 3:275–288.

Keinan, A., Sandbank, B., Hilgetag, C., Meilijson, I., and Ruppin, E. (2004). Fair attribution of functional contribution in artificial and biological networks. *Neural Computation*, 9(16):1887–1915.

Kendal, J., Ihara, Y., and Feldman, M. (2005). Cultural niche construction with application to fertility control: A model for education and social transmission of contraceptive use. *Morrison Institute for Population and Resource Studies, Working Paper*, 102.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Kirby, S. and Hurford, J. (1997). Learning, culture and evolution in the origin of linguistic constraints. In Husbands, P. and Harvey, I., editors, *4th European Conference on Artificial Life*, volume IV, pages 493–502, Cambridge, MA. MIT Press.

Kirkpatrick, S., Gelatt, C., and Vecchi, M. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

Korona, R., Nakatsu, C., Forney, L., and Lenski, R. (1994). Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proceedings of the National Academy of Sciences USA*, 91:9037–9041.

Kozima, H., Nakagawa, C., and Yano, H. (2002). Emergence of imitation mediated by objects. In Prince, C., Demiris, Y., Marom, Y., Kozima, H., and Balkenius, C., editors, *Proceedings Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, volume 94, pages 59–61, Edinburgh, Scotland.

Kussell, E. (2005). The designability hypothesis and protein evolution. *Protein and Peptide Letters*, 12(2):111–116.

Laland, K. (1992). A theoretical investigation of the role of social transmission in evolution. *Ethology and Sociobiology*, 13(2):87–113.

Laland, K., Odling-Smee, F., and Feldman, M. (1996). The evolutionary consequences of niche construction: A theoretical investigation using two-locus theory. *Journal of Evolutionary Biology*, 9:293–316.

Laland, K., Odling-Smee, F., and Feldman, M. (2001). Cultural niche construction and human evolution. *Journal of Evolutionary Biology*, 14:22–33.

Langton, C. (1988). Artificial life. In Langton, C., editor, *Artificial Life, volume VI of Santa Fe Institute Studies in the Sciences of Complexity*, pages 1–47. Addison-Wesley, Reading, MA.

Langton, C. (1995). *Artificial Life: An Introduction*. MIT Press, Boston, MA.

Lenski, R., Ofria, C., Collier, T., and Adami, C. (1999). Genome complexity, robustness and genetic interactions in digital organisms. *Nature*, 400(6745):661–664.

Lewontin, R. (1983). Gene, organism, and environment. In Bendall, D., editor, *Evolution from Molecules to Men*. Cambridge University Press.

Littman, M. (1996). Simulations combining evolution and learning. In Belew, R. K. and Mitchell, M., editors, *Adaptive Individuals in Evolving Populations: Models and Algorithms*, pages 465–477. Addison Wesley, Reading, MA.

Littman, M. and Ackley, D. (1991). Adaptation in constant utility non-stationary environments. In Belew, R. and Booker, L., editors, *Proc. of the Fourth Int. Conf. on Genetic Algorithms*, pages 136–142, San Mateo, CA. Morgan Kaufmann.

Lynch, A. (1999). *Thought Contagion: How Belief Spreads Through Society*. Basic Books, New York, NY.

Mace, R. (1996). When to have another baby: A dynamic model of reproductive decision-making and evidence from gabbra pastoralists. *Ethology and Sociobiology*, 17:263–273.

Mace, R. (2000). An adaptive model of human reproductive rate where wealth is inherited: Why people have small families. In Cronk, L., Chagnon, N., and Irons, W., editors, *Adaptation and Human Behaviour*, pages 261–282. Aldine de Gruyter, New York, NY.

Macken, C. and Perelson, A. (1989). Protein evolution on rugged landscapes. *Proceedings of the National Academy of Sciences USA*, 86(16):6191–6195.

Marom, Y., Maistros, G., and Hayes, G. (2002). Toward a mirror system for the development of socially-mediated skills. In Prince, C., Demiris, Y., Marom, Y., Kozima, H., and Balkenius, C., editors, *Proceedings Second International*

Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems, volume 94, Edinburgh, Scotland.

Mayley, G. (1996). Landscapes, learning costs and genetic assimilation. *Evolutionary Computation*, 4(3):213–234.

Mayley, G. (1997). Guiding or hiding: Explorations into the effects of learning on the rate of evolution. In Husbands, P. and Harvey, I., editors, *Proceedings of the Fourth European Conference on Artificial Life*, pages 135–144, Cambridge, MA. Bradford Books/MIT Press.

Maynard-Smith, J. (1987). Natural selection: When learning guides evolution. *Nature*, 329:761–762.

Meilijson, I. (2003). The time to a given drawdown in Brownian motion. *Séminaire de Probabilités XXXVII., Springer Lecture Notes in Mathematics*, 1832:94–108.

Meltzoff, A. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, 24:470–476.

Meltzoff, A. (1996). The human infant as imitative generalist: A 20-year progress report on infant imitation with implications for comparative psychology. In Heyes, C. and Galef, B., editors, *Social Learning in Animals; The Roots of Culture*, New York Academic Press.

Meltzoff, A. (2002). Elements of a developmental theory of imitation. In Meltzoff, A. and Prinz, W., editors, *The Imitative Mind: Development, Evolution and Brain Bases*, pages 19–41. Cambridge University Press, Cambridge, MA.

Meltzoff, A. and Moore, K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6:179–192.

Menczer, F. and Belew, R. (1994). Evolving sensors in environments of controlled complexity. In Brooks, R. and Maes, P., editors, *Artificial Life IV*, pages 210–221, Cambridge, MA. MIT Press.

Mery, F. and Kawecki, T. (2002). Experimental evolution of learning ability in fruit flies. *Proceedings of the National Academy of Sciences USA*, 99:14274–14279.

Mery, F. and Kawecki, T. (2004). The effect of learning on experimental evolution of resource preference in *drosophila melanogaster*. *Evolution*, 58:757–767.

Mills, R. and Watson, R. (2006). On crossing fitness valleys with the baldwin effect. In Rocha, L., Bedau, M., Floreano, D., Goldstone, R., Vespignani, A., and Yaeger, L., editors, *Proceedings of Tenth International Conference on the Simulation and Synthesis of Living Systems (ALifeX)*, pages 493–499.

Mitchell, M. (1998). *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA.

Morgan, C. (1896). On modification and variation. *Science*, 4:733–740.

Moriarty, D. and Mikkulainen, R. (1996). Efficient reinforcement learning through symbiotic evolution. *Machine Learning*, 22(1-3):11–32.

Moritz, E. (1990). Memetic science: I - General introduction. *Journal of Ideas*, 1:1–23.

Mühlenbein, H., Schomish, M., and Born, J. (1991). The parallel genetic algorithm as function optimizer. *Parallel Computing*, 17:619–632.

Munroe, S. and Cangelosi, A. (2002). Learning and the evolution of language: The role of cultural variation and learning costs in the baldwin effect. *Artificial Life*, 8(4):311–339.

Nehaniv, C. and Dautenhahn, K. (2002). The correspondence problem. In Dautenhahn, K. and Nehaniv, C., editors, *Imitation in Animals and Artifacts*. The MIT Press.

Nielsen, R. and Yang, Z. (2003). Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Molecular Biology and Evolution*, 20(8):1231–1239.

Nolfi, S., Elman, J., and Parisi, D. (1994). Learning and evolution in neural networks. *Adaptive Behavior*, 1(3):5–28.

Nolfi, S. and Floreano, D. (1999). Learning and evolution. *Autonomous Robots*, 7(1):89–113.

Nolfi, S. and Parisi, D. (1997). Learning to adapt to changing environment in evolving neural networks. *Adaptive Behavior*, 1:99–105.

Noskowicz, S. and Goldhirsch, I. (1990). First passage time distribution in random random walk. *Physical Review A*, 42:2047–2064.

Notestein, F. (1953). Economic problems of population change. In *Proceedings of the Eighth International Conference of Agricultural Economists*, pages 13–31, London. Oxford Univ. Press.

Odling-Smee, F., Laland, K., and Feldman, M. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton University Press, Princeton.

Oksanen, L. and Lundberg, P. (1995). Optimization of reproductive effort and foraging time in mammals: The influence of resource level and predation risk. *Evolutionary Ecology*, 9(1):45–56.

Oztop, E. and Arbib, M. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, 87:116–140.

Paenke, I., Kawecki, T., and Sendhoff, B. (2006). On the influence of lifetime learning on selection pressure. In Rocha, L., Bedau, M., Floreano, D., Goldstone, R., Vespignani, A., and Yaeger, L., editors, *Proceedings of Tenth International Conference on the Simulation and Synthesis of Living Systems (ALifeX)*, pages 500–506.

Parisi, D. and Nolfi, S. (1996). The influence of learning on evolution. In Belew, R. K. and Mitchell, M., editors, *Adaptive Individuals in Evolving Populations: Models and Algorithms*, pages 419–428. Addison Wesley, Reading, MA.

Pigliucci, M. (2001). *Phenotypic Plasticity: Beyond Nature and Nurture*. The Johns Hopkins University Press, Baltimore, MD.

Prinz, W. and Meltzoff, A. (2002). An introduction to the imitative mind and brain. In Meltzoff, A. and Prinz, W., editors, *The Imitative Mind: Development, Evolution and Brain Bases*, pages 1–15. Cambridge University Press, Cambridge, MA.

Ramnani, N. and Miall, R. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience*, 7:85 – 90.

Rizzolatti, G. and Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences*, 21:188–194.

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (2002). From mirror neurons to imitation: Facts and speculations. In Meltzoff, A. and Prinz, W., editors, *The Imitative Mind: Development, Evolution and Brain Bases*, pages 247–266. Cambridge University Press, Cambridge, MA.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3:131–141.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2:661–670.

Rogers, E. M. (1962). *Diffusion of Innovations*. Free Press, New York, NY.

Ruppin, E. (2002). Evolutionary autonomous agents: A neuroscience perspective. *Nature Reviews Neuroscience*, 3(2):132–141.

Salomon, R. (1996). Re-evaluating genetic algorithm performance under coordinate rotation of benchmark functions: A survey of some theoretical and practical aspects of genetic algorithms. *BioSystems*, 39:263–278.

Schlichting, C. and Pigliucci, M. (1998). *Phenotypic Evolution: A Reaction Norm Perspective*. Sinauer Associates Inc., Sunderland, MA.

Schuster, P., Fontana, W., Stadler, P., and Hofacker, I. (1994). From sequences to shapes and back: A case study in RNA secondary structures. *Proceedings of The Royal Society, Biological sciences*, 255:279–284.

Spitzer, F. (2001). *Principles of Random Walk*. Springer, New York, NY, 2nd edition.

Stadler, P. (1995). Towards a theory of landscapes. In Lopéz-Peña, R., Capovilla, R., García-Pelayo, R., Waelbroeck, H., and Zertuche, F., editors, *Complex Systems and Binary Networks*, pages 77–163. Springer Verlag, Berlin, New York.

Todd, P. and Miller, G. (1991). Exploring adaptive agency II: Simulating the evolution of associative learning. In Meyer, J. and Wilson, S., editors, *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pages 306–315, Cambridge, MA. MIT Press.

van Nimwegen, E., Crutchfield, J., and Huynen, M. (1999). Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences USA*, 96:9716–9720.

Waddington, C. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150:563–565.

Waddington, C. (1953). Genetic assimilation for acquired character. *Evolution*, 7:118–126.

Wahlund, S. (1928). Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungs-lehre ausbetrachtet. *Hereditas*, 11:65–106.

Weber, B. and Depew, D., editors (2003). *Learning and Evolution. The Baldwin Effect Reconsidered*. MIT Press, Cambridge, MA.

Weinberger, E. (1990). Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biological Cybernetics*, 63:325–336.

West-Eberhard, M. (2003). *Developmental Plasticity and Evolution*. Oxford University Press, New York, NY.

Whiten, A., Goodall, J., McGrew, W., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C., Wrangham, R., and Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399:682–685.

Whiten, A. and Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. In Slater, P., Rosenblatt, J., Beer, C., and Milinski, M., editors, *Advances in the Study of Behavior*, pages 239–283, San Diego, CA. Academic Press.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In Jones, D., editor, *Proceedings of the Sixth International Congress of Genetics*, volume 1, pages 356–366, Brooklyn Botanic Garden, Brooklyn, New York, NY.

Zentall, T. (2001). Imitation in animals: Evidence, function, and mechanisms. *Cybernetics and Systems*, 32(1-2):53–96.