

ABSTRACT OF THE DISSERTATION

The multiplicative weight updates method for evolutionary biology

by Erick Chastain

Dissertation Director: Eric Allender

A new and exciting direction of recent work in theoretical computer science is the application of methods from the field to evolutionary biology. Starting with the work of Christos Papadimitriou and Adi Livnat, there has been a concerted effort to use these techniques to analyze such diverse phenomena as: the algorithmic role of recombination to increase mixability, the evolution of modularity, and the evolution of complex adaptations. There is also work by Les Valiant and his students using tools primarily from learning theory to more broadly analyze evolutionary processes. In parallel, computer science theory has developed a novel method which has been applied in diverse areas of algorithms and complexity: the Multiplicative Weight Updates (MWU) method. The MWU method simply applies the MWU general-purpose online learning algorithm on problem-specific loss functions. The contribution of this thesis is to apply the MWU method and the algorithmic lens to make models in evolutionary biology.

The first contribution is a surprising equivalence between the MWU algorithm playing a coordination game and infinite-population genetics models with recombination and no mutation. By so doing, we resolve analytically a question asked by Papadimitriou and Livnat: whether mixability is increased in the short-term by recombination.

Other models introduced using MWU as a basic dynamics include a model of the

evolution of animal personality and of tool innovation.

Finally, the thesis presents a novel connection between universal semantic communication and the Rivoire-Leibler model of population genetics, in addition to infinite population asexual selection models. MWU is the basic tool used to prove the latter connection.

Acknowledgements

Thanks to my advisor Eric for his tireless support and always watching out for me. Thanks to Nina Fefferman for helping me with much of the preparatory work on this thesis, and helping along the way. Finally thank you to rest of my committee for all their help, especially Shubhangi who was also on my quals committee and was always encouraging me to investigate interesting ideas.

Dedication

To Lucy, my light. To St. Augustine, my Oarsman. And to Rosemary Chastain, who brought me back to shore, *Requiem aeternam dona ei, Domine, et lux perpetua luceat ei. Requiescat in pace.*

Table of Contents

Abstract	ii
Acknowledgements	iv
Dedication	v
1. Introduction	1
1.1. Evolution as an Algorithm, or the use of Algorithmic methods in Evolutionary Biology	2
2. Evolutionary biology and Population Genetics	10
2.1. Biological Evolution	10
2.1.1. Natural selection	15
Principles of Natural Selection	17
2.1.2. Variation and Population Thinking	20
2.2. Population Genetics	24
2.3. Mixability	30
3. The Multiplicative Weight Updates method	32
3.1. Multiplicative Weight Updates	32
3.1.1. MWUA as a Game dynamics	37
3.1.2. The MWU Method and its Applications	38
4. Population Genetics and MWUA	42
4.1. MWUA and Population Genetics	42
4.2. Supplementary Information	51
4.3. Population Genetics under selection and recombination	51
4.4. The Experts Problem	54

4.5.	An Optimization Interpretation of MWUA	55
4.6.	Coordination Games between Genes	56
4.6.1.	Generalizations: Strong selection and Diploid organisms	64
5.	Evolution of Animal Personality	73
5.1.	Supplemental Information	82
6.	Evolution of Tool Innovation	86
6.1.	Introduction	86
6.2.	Model	90
6.2.1.	Description and Assumptions	90
6.2.2.	Formal description	91
6.3.	Discussion	93
6.4.	Conclusion	95
6.5.	Supplementary Information	95
7.	Process Information and Evolution	99
7.1.	Introduction	99
7.2.	Process information	103
7.3.	Process information and universal semantic communication	107
7.4.	Online learning, Multiplicative Weight Updates, and Evolution	109
7.5.	Process Information and Evolution	111
7.6.	Proofs of Evolutionary theory and MWU Results	119
7.7.	Proofs of Process information and USC Results	124
7.8.	Category Theory	127
7.9.	Category-Theoretic Process Information	128
8.	Conclusion	130
8.1.	Population Genetics and MWU	130
8.2.	Evolution of Animal Traits	134
8.3.	Over-arching themes	135

References 137

Chapter 1

Introduction

One of the most enduring mysteries in science is the origin of the manifold forms of life. In the words of Charles Darwin:

“There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.”

All of the diversity of life is accompanied by complex and intricate structures. A great stimulus for the life sciences is how such complex structures could possibly have originated. There have been many theories of the origin of complex structures in living things, but most of them have focused on some kind of gradual process of development or evolution. Evolution would operate on a source of variation, and then originate new forms that gradually increase the rate of reproduction and survival, called fitness. There are of course many remaining questions, such as how it could be possible for novel forms to originate with small changes. And moreover, how the sheer vastness of different forms could have emerged. In particular, in population genetics (PG), there is a population of different kinds of organisms, and the frequency of the kind with the highest fitness will tend to increase.

Other mysteries have emerged as well, such as, how could it be that sexual reproduction ever got going, given that most living organisms started off as asexual, and asexual reproduction is the most efficient way to reproduce. Finally, how could the forms of life occurring in history originate, in other words if one rewound the tape of history and started it over again, would the same or similar forms of life have come about? This mystery is referred to as the “arrival of the fittest,” a play on the “survival

of the fittest.”

Why as a computer scientist am I interested in Evolution? Computer science looks at algorithmic processes that use and quantify resources such as time and randomness, which are then levied for the solving of problems. In particular, as computer scientists we can quantify how much time and how much randomness (or some other resource) are necessary to solve problems of increasing problem size. So therefore with our tools we can study the question: what process could have generated all this diversity and complexity in so few generations? Also, how can we shed light on the question of the origin of sexual reproduction? Does it really make sense in a broad sense to shuffle genetic material in order to maximize fitness? If not, what is maximized? What kinds of algorithms can be so versatile as to solve many different problems, if biological structures are each formalized as “problems?”

There has been much interest in evolutionary biology from computer scientists in algorithms and complexity theory, for exactly these reasons. Applying rigorous tools from theoretical computer science, the community has produced a body of work in this exciting new interdisciplinary field. First we will discuss the different subdisciplines of computer science and their interaction with evolutionary biology: learning theory, optimization theory, and algorithmic analysis.

1.1 Evolution as an Algorithm, or the use of Algorithmic methods in Evolutionary Biology

The interaction between evolutionary theory and algorithmic thinking is almost as old as computation itself, with Charles Babbage imagining the origin of species as progressing according to some gradual algorithm [38]. John Von Neumann too tinkered with modeling mechanisms in evolution computationally, by thinking up a machine that could reproduce itself [139]. In the process he invented cellular automata. There have been many philosophers or evolutionary biologists who have toyed with the idea of an evolutionary process being algorithmic in some way, including Richard Dawkins

with his WEASEL program [36], Daniel Dennett with his idea of a “Darwinian algorithm,” [38], and John Maynard Smith applying a kind of dynamical game theory to the study of evolution and advancing the idea that evolution operates on, and innovates, information-processing structures [91, 132]. Most notably, the work of William Fontana actually uses a Lambda-calculus formalism to analyze how evolution could produce innovative new types that replicate themselves [48].

Among all of the evolutionary biology approaches to information-processing, most of them focus very keenly on the mechanistic nature of natural selection, the means by which fitness is increased in evolving populations. However, in addition, they have noticed that evolution itself operates on discrete structures which have a kind of computational logic, and which process information. If one tries to modify algorithms by strange random additions, one gathers a certain intuition about which algorithms are robust to changes, and which ones are brittle. These studies have culminated in the analysis of robustness and evolvability [141, 76]. Also, the very structure of organisms has been analyzed along the lines of divide-and-conquer algorithms, via modular theories of organismal structure [142]. There has also been much work making biologically-inspired algorithms for optimization, such as genetic algorithms.

But within theoretical computer science itself, as an outgrowth of previous approaches and intuitions, there has been a novel emphasis on growing the seeds planted by these researchers. The first work to do this in earnest was done by Les Valiant and his students [137]. The basic idea he had was to view evolution as a kind of learning algorithm, trying to use noisy samples of a quality function to judge whether an individual hypothesis that was being modified slowly over evolutionary time was of increasing quality. If so, then the hypothesis would be retained. And so forth. The interesting thing about the model was his emphasis on the individual (which is called strong-selection weak mutation in population genetics), and the vast and robust model of evolution assumed: almost any mechanism that has been imagined could fit into the model of evolution proffered. Through this model, it was shown that certain classes of functions could not ever be realized by evolution in realistic time-scales, such as the XOR function. Moreover, evolution was shown to be as powerful as a relatively weak

form of learning called statistical query learning, which is less powerful than conventional machine learning.

In another area of theoretical computer science, algorithmic analysis, Bernard Chazelle and co-authors have focused on “natural algorithms.” Natural algorithms are algorithms used to model biological function, that are shown to be reasonably optimal by some performance measures. It is often supposed that Evolution optimizes the function of such algorithms in some way (which need not be the case). In evolutionary theory, this fits in trait-based evolutionary biology. The field has seen progress in analyzing the amount of time a famous model of bird-flocking takes to converge [27], in addition to another model of biological synchronization [26].

The early history of evolutionary computation was to think of genetics as a means to optimize some objective function [60]. In that context, strong-selection weak-mutation regimes (with recombination) for population genetics do give the kind of algorithm they implement. Such an approach seeks to use population genetics to enable global optimization. There has been much work in theoretical computer science trying to understand when, if at all, such a heuristic would work. There have been mixed results [7, 3].

However there has recently been work by Christos Papadimitriou and co-authors that has sought to find conditions under which evolution with recombination could be useful, and others in which it is not so useful [84]. Papadimitriou et. al. do this by asking whether another quantity, “mixability” is optimized by evolution with recombination, rather than fitness itself. Mixability is basically a quantity that measures how well a gene performs in a different set of genetic contexts. And finally, note that there are some problems (specifically, finding satisfying assignments) for which they have shown that recombination + selection suffices [87].

The general trend in all of these applications of computational complexity and algorithms is to find the suitable problem that Evolution putatively solves, then show that it can be used as an algorithm for that problem. For instance, Valiant’s Evolvability framework looks at learning problems based on noisy queries, and the work on mixability looks at optimization. For this thesis, we focus fundamentally on a decision-theoretic

problem, in a setting called on-line learning. We want to decide between different experts and arbitrate between them based on the quality l_i of the advice they give—yielding choices that are close-to-optimal when compared to the best expert in hindsight. The relation between this basic problem of expert-based learning and Evolution is what we will study in various parts of the thesis. We also look at a special case in which the experts are pure strategies in a game (Chapter 3).

The principle algorithm we propose to use in modeling in evolutionary biology is called the Multiplicative Weight Updates Algorithm (MWUA). The basic principle behind the MWUA is to trust the advice of an expert with a probability proportionate to the quality of advice, according to a simple product of the current probability and a factor directly proportional to the advice quality.

There are many useful properties of MWUA that we use throughout the thesis:

- MWUA works for any set of quality functions, even those that are adversarially chosen.
- The algorithm chooses a set of probabilities over experts that does as well cumulatively as the best expert.
- The algorithm converges very quickly as a function of the total elapsed time T , often of order $1/T$.
- Because of the exponential/linear form of the probability update, MWUA is “natural,” or “intuitive.” Therefore, many dynamics/algorithms/gadgets can be analyzed as a variant of MWUA.

Given all of these properties, the MWUA is very useful for solving a variety of optimization problems. The application of MWUA to solve an optimization problem is called the Multiplicative Weight Updates (MWU) method. Specifically, the MWU method (as introduced by Kale [6]) works as follows:

Apply MWUA with experts and l_i chosen differently based on the problem class, for optimization/learning problems. The final mixture over experts must then be used to give an approximately-optimal solution to the given problem.

There have been applications so far to solving the following kinds of problems:

- Solving zero-sum games approximately in quadratic time
- Linear program solver
- Convex program solver
- Learning linear classifiers
- Boosting

Now after showing that the MWU method has had great success in solving a variety of problems, we show that it is also useful for models in Evolutionary biology. Many of the tools used in applications of the MWU method to the above problems are used again to inspire and prove properties of the models.

The principal contributions of this thesis are to show just how versatile the MWU method is for generating novel models and exciting connections between different fields of evolutionary biology and computer science theory. In particular, we want to introduce the MWU as an algorithm to generate dynamics for use in biological modeling.

The general-purpose description for how we can use the MWU method in biological models centers on biological information. First of all, choose the receivers of biological information for the target domain as experts. Then choose the appropriate quality function l_i proportionate to the value of information for expert i . We will call this the Expert Information Model (EIM).

Now let's be more specific about the particular results proven in the thesis using the EIM, and thus the MWU method.

Here are examples of some things we prove in the thesis using this method, along with a description of each result:

- Population genetics as MWUA playing a Coordination game (w/ Papadimitriou, Vazirani, Livnat) (Chapter 4)

Even the most seasoned students of evolution, starting with Darwin himself, have occasionally expressed amazement that the mechanism of natural selection has

produced the whole of Life as we see it around us. There is a computational way to articulate the same amazement: “What algorithm could possibly achieve all this in a mere three and a half billion years?” In this work we propose an answer: We demonstrate that in the regime of weak selection, the standard equations of population genetics describing natural selection in the presence of sex become identical to those of a repeated game between genes played according to multiplicative weight updates (MWUA), an algorithm known in computer science to be surprisingly powerful and versatile. MWUA maximizes a trade-off between cumulative performance and entropy, which suggests a new view on the maintenance of diversity in evolution.

According to mixability theory, in the presence of sex and natural selection, alleles are favored that perform well across a wide variety of genetic combinations. This theory was formed as a bridge between evolutionary biology and computer science, informing both evolutionary theory and evolutionary computation. Recently, the connection between computer science and evolutionary biology has been extended. It was shown that, in the realm of weak selection, using a haploid model, the population genetic dynamics in the presence of natural selection and sex are equivalent to a repeated coordination game between genes, where the loci are the players, the alleles are the strategies, fitness is utility, and the game is played according to a surprisingly powerful algorithm known in computer science as the Multiplicative Weights Update Algorithm (MWUA), and known in other fields by names including softmax, exponential-family Bayesian updating, and the Boltzmann distribution. This connection between MWUA and game-theory enables the application of the MWUA no-regret theorem to the population genetic context. Here, we bring this application to the attention of theoretical biologists, and provide one interpretation of it in terms of mixability: the cumulative population mean fitness is close to the cumulative mixability of the most mixable alleles. We show that this holds for haploid models with both weak and strong selection, and for diploid models, extending our analytical understanding of mixability and suggesting that mixability is an important guide to population genetic dynamics.

- Link between Universal Semantic Communication [67] and Population genetics by using EIM (w/ Cameron Smith) (Chapter 7)

Juba and Sudan's Universal Semantic Communication (USC)[67, 68] is a theory that models communication among agents without the assumption of a fixed protocol. We demonstrate a connection, via process information, between a special case of USC and evolutionary processes. In this context, one agent attempts to interpret a potentially arbitrary signal produced within its environment. Sources of this effective signal can be modeled as a single alternative agent. Given a set of common underlying concepts that may be symbolized differently by different sources in the environment, any given entity must be able to correlate intrinsic concepts with input it receives from the environment in order to accurately interpret the ambient signal and ultimately coordinate its own actions. This scenario encapsulates a class of USC problems that provides insight into the semantic aspect of a model of evolution proposed by Rivoire and Leibler [120]. Through this connection, we show that evolution corresponds to a means of solving a special class of USC problems, is a special case of the Multiplicative Weights Updates algorithm, and that infinite population selection with no mutation and no recombination conforms to the Rivoire-Leibler model. Finally, using process information we show that evolving populations implicitly internalize semantic information about their respective environments.

- Evolution of Animal Personality using the EIM, with the MWUA as the base dynamics (w/ Nina Fefferman) (Chapter 5)

Animal personality traits and the costs and benefits from those traits in different contexts are of contemporary interest, with many recent studies exploring these topics. In this chapter we propose a general framework to understand the evolution of personality itself (rather than the individual traits, or small suites of traits). The framework encompasses the different theories of animal personality traits so far proposed. We apply the MWU method to study the evolutionary

benefits of having a 'personality,' and apply this to cases of interest. Each personality is a probability over different complex personality bundles. Lastly, we use this framework to discuss how insights from previous work can, together, describe the fitness landscape of personalities in different types of environments.

- Evolution of tool-creation using the EIM, with an infinite-expert variant of MWUA as basic dynamics (w/ Nina Fefferman) (Chapter 6)

Tool innovation by animals is a truly remarkable phenomenon in which genuinely novel tools are produced from raw materials, seemingly from nothing. Novel work in New Caledonian crows has shown that tools can be created without any similar tool previously existing. The evolutionary basis for the development of tool use is still uncertain, with few theories extant to explain it. We advance a novel theory of tool innovation and its evolution based on a theory of novelty that draws a tool from an infinite pool of possible tools. A crucial component in this model is a novel variant of the MWUA, which draws experts from a distribution with infinite support. We show how our novel theory unifies previously published hypotheses and predicts some of the otherwise unexplained observations of tool use, including regression to caching behavior in rooks.

We will first proceed to give some background information on Evolutionary biology in Chapter 2, then proceed to describe the MWU method in Chapter 3. Then, we outline the applications of the MWU method to Population genetics (Chapter 4), Evolution of Animal Personality (Chapter 5), Evolution of Tool-innovation (Chapter 6), and Evolution as a population acquiring semantic information (Chapter 7). Finally we will finish with discussion of interesting related issues and follow-up work in the literature based on our work, and also future work (Chapter 8).

Chapter 2

Evolutionary biology and Population Genetics

This chapter describes the fundamentals of Evolutionary theory and Population genetics, with a keen focus on areas of interest for the thesis. In addition to giving a general over-view of biological evolution (in Section 2.1) and how it works, population genetics terms and equations are given (in Section 2.2) to facilitate reading of the technical parts of the thesis devoted to population genetics. Note that we do not claim any completeness in the survey of material we present here, and interested readers should consult appropriate references [9, 14, 102] for more details.

2.1 Biological Evolution

In *On the Origin of Species*, Darwin proposed two big ideas, which had been proposed before, but never with as much evidence or mechanism as he marshaled to the task. After collecting samples from around the world, either himself or from naturalists abroad, and trying to synthesize extant theories, Darwin recognized that the current diversity of life descends from previously existing species. A common ancestor is posited that all such species descend from. Secondly, the fit of species to their environment is the result of **natural selection**, a gradual process by which forms that are better suited to the environment increase in frequency in a population. Another way to understand natural selection is as a kind of “sieve” for a population with different forms of life, with the environment sifting out varieties that are not well-adapted to it. The time that the population under selection takes to boost the better suited forms ranges from days to tens of thousands of years, depending on a number of factors, including the strength of selection, and the reproduction rate of the organisms under study. As a whole, the two major ideas proposed by Darwin suggest that all of the forms of life around us are

ever-developing and dynamic. The impact of Evolutionary biology on subfields of biology is immense, so much so that renowned biologist Theodosius Dobzhansky remarked “nothing in biology makes sense except in the light of evolution.” [41]

There are many foundational issues to discuss in Darwin’s account, some of which have recurred in recent years in both the complex systems and theoretical biology communities. We will return to these later. Keeping this in mind, we proceed in discussing the exciting science of evolutionary biology. Evolution seeks to find the root causes of and relationships between all forms of life, as seen through the lens of their many different manifestations in genetic material, fossil records, morphology, and behavior. One could say that evolution is like an artist, introducing with every stroke a new variant of an underlying way of living, all unified by a graspable form. The commonalities in way of life introduced one can discover via experimentation and observation. The kinds of data typically used to analyze such things include anatomical, molecular, genetic, developmental, and behavioral data; and more specifically, DNA sequences, population composition, and imaging of relevant anatomy. Older approaches to imaging, such as that taken by von Humboldt and Haeckel [138, 55] focused on detailed anatomical sketches in the natural habitat. And the scoring system to analyze population composition focused on different apparent traits of the organisms, and counting their frequencies in the population. We will return to these traits in the sequel.

Evolutionary biology is the study of the origin, persistence, and manifold variety of life on Earth for the last 3.5 billion years. A species evolves from its ancestral species, from which it *descended* and *modifications* have occurred. Darwin called this whole process **descent with modification**. For example, if we want to understand the evolution of *Homo sapiens*, we need to find the ancestral species of primate from which it descends, and the modifications that happened along the way.

Among causes of modification, the most important is natural selection. We simply outline briefly here a basic sketch of how selection operates, and leave a more detailed discussion for later. Genetic **mutations**, or changes in the DNA sequence, arise rarely, but continually, due to environmental factors, copying errors, or other unknown causes.

The mutations, which change the **genotype** at particular sites (**loci**) of the DNA sequence, in turn change the **phenotype** — the observable, graspable traits (**alleles**) — of organisms. Note that the characteristic phenotype of an organism need not be any necessary trait to understand that species apart from others, but rather it is a “tell” or an accidental feature that distinguishes it from the rest. These mutations can reduce, increase, or keep unchanged the fitness, in which case they are called deleterious, beneficial, or neutral (respectively) — where fitness is measured using rates of reproduction and relative survival. It is quite common for deleterious mutations to arise in very finely-tuned organisms. To see why, by analogy consider an out-of-place wire in a wiring diagram for building a circuit board, and its effect on the function of the manufactured chip. However, it is known that, unlike circuit boards, for things like metabolism, most often mutations are neutral or even beneficial, increasing energetic efficiency [122]. So in many ways the effect on fitness of a mutation depends on the phenotype it is linked to. If a beneficial mutation arises by chance, then it will increase the rates of reproduction or survival, and thereby increase the frequency of variants with that mutation. Such a result is called evolutionary change by natural selection.

The beneficial genetic changes, accumulated over many generations, can produce quite drastic effects within a population. Even small changes, as shown by [122], can have a lasting effect on the phenotype. The effects either way of these new accumulated mutations can produce new species (by a process called speciation), genera, families, and higher taxonomic orders. One proposed mechanism for speciation has focused on geographic or reproductive isolation due to migration or other factors, in order to give the population time to diverge from its original population and accumulate the necessary mutations [92]. Sometimes, in fact, sets of beneficial mutations have led to major transitions in biological history, such as the evolution of prokaryotes, eukaryotes, multicellular organisms, et cetera.

The most elementary examples of evolution, and indeed some most inspiring to Darwin, were based on selective breeding over time. Selective breeding has happened throughout history for the purpose of food cultivation, with such crops as barley (*Hordeum vulgare*), wheat (*Triticum*), as well as lentils (*Lens culinaris*), and peas

(*Pisum sativum*), for over 10,000 years [53, 152, 1].

In order to understand natural selection, we must first understand selective breeding. The process of selective breeding by humans, known as artificial selection, is easy to understand, and thus a good starting point for further examples. For crops, in each generation the best (in size, survival rate, and flavor) plants are chosen as parents for the next generation. Artificial selection is related to natural selection. With natural selection, traits associated with increased survival and reproduction — traits compatible with flourishing of the organism — increase in frequency. In the case of artificial selection, traits compatible with human flourishing are selected to increase in frequency. Food cultivation has produced amazing effects, such as doubling the yields of wheat, rice, and corn since 1930 [64, 110, 42].

However, sometimes human ingenuity paired with natural selection can work towards unintended consequences that are detrimental to human flourishing. Such is the case for crop yields. By the use of pesticides, we have been able to increase crop yields substantially by killing off pests that eat crops. However, because of the much-faster reproduction rates of the insects, they are able to develop resistance to the pesticide, and we must continually innovate new pesticides to keep the boost in crop yields that pesticides afford. How can we say that this is natural selection, instead of artificial selection, given that humans are using and distributing pesticides? The difference between artificial selection and natural selection is that the latter involves no deliberate human choices as to who will reproduce, not whether humans are involved in the process at all.

All living things have descended from a common ancestor, and over generations ancestors diversified more and more to yield the impressive variety of life forms we observe today. Darwin pioneered a way of thinking centered on this conception, by use of “phylogeny” and rooted trees that visualize descent from the common ancestors. The parents on such trees are common ancestors, and the children are descendants. All species living on the Earth in the present or the past are placed on the “tree of life.” Such trees are also called “phylogenetic trees.” Leaves of the tree of life are species currently living today or which have gone extinct. Branching points in the tree are “divergence

events”— events which cause speciation or the arrival of a new kind of organism— which happen in the past. We have discussed diversity very informally, but now we can describe it in more detail: phylogenetic diversity is the number of children in a phylogenetic tree— the number of distinct lineages that are descendants of a particular species. If one cares about conservation of biodiversity, then one would perhaps want to set environmental policies that are mindful of how rapidly speciation occurs in a given branch of the phylogenetic tree. Why would we care about biodiversity? An ecosystem can be more resilient to environmental changes if there is abundant biodiversity. As such, not only are there scientific merits to understanding the maintenance of diversity, but also applications to the “care for our common home” [50].

As a field, evolutionary biology has included both experimental and theoretical approaches. The latter are not well-known outside of the field. The former include some of the first results: including the finding of a relationship between dog-like species like *pakicetus* and whales, the surprising similarity anatomically between humans and other primates (stemming from a predicted common ancestor) including such shared features as the intermaxillary bone [54], and the finding that there is much shared in the respective DNA sequences that bears out these similarities.

The role of theory in evolutionary biology is to shape and further the research agenda for the field by analyzing how complicated systems work, allowing for predictions and inferences and synthesizing previous theories (as Darwin did). Theoretical biology often uses mathematical models, especially for complex biological systems. A good model distills the essence of the biological system, focusing on only the most critical components. The most common use of models is to plan for the future and make predictions. In evolutionary biology, this is especially useful for conservation and ecology. For instance, when ecologists try to design captive breeding plants, they use population genetics (which we will introduce in the sequel) to ensure there is enough diversity in the resulting population. Other applications of models include the use of models for neutral mutation to predict how long a DNA sequence takes to change, which has its use in reconstructing phylogenetic trees. Finally, not all models in evolutionary biology are used to predict how evolution itself changes the frequency of different

phenotypes, but rather how biological systems could have evolved to begin with. For example, a model of how a combinatorial alphabet evolved for humans but not for primates [78]. The two families of evolutionary models we will study in this thesis are outlined broadly above:

1. Population genetics: the study of how phenotype frequencies change in a population under natural selection.
2. Arrival of the Fittest [48], how did the fittest originate?: the study of how a given biological system could have evolved, or have arisen as a possible phenotype. This thesis will focus on the former.

2.1.1 Natural selection

As mentioned in the more general introduction of evolutionary biology, artificial selection was an important building block for Darwin’s theory of natural selection. Recall that Darwin’s book introducing natural selection was describing the origin of species. The contemporaries of Darwin accepted that artificial selection could produce new varieties of pigeons— new behaviors, new feather color/patterning, et cetera. But the idea that artificial selection could create new species was entirely unheard-of, as then it raised the possibility of a new and original life-form emerging. Therefore, Darwin spent much time arguing that species are “strongly marked and permanent varieties.” Using many examples from selective breeding in animals and plants, Darwin was able to show that new varieties could arise from a single stock of organisms— using precursors of phylogenetic trees, with descent and modification. Then, having established that certain kinds of varieties are species, Darwin could claim that similar processes generated both. Thus he could claim that a kind of selection process could operate to produce both varieties and species, and therefore that new species arise from other species. Having established that artificial selection could possibly lead to changes in species, Darwin introduces by analogy a new kind of selection process, natural selection, as what happens to generate natural species in the absence of human selection.

For natural selection, the selective agent is nature (as compared to the human for

artificial selection), but one should not understand nature as being a conscious agent in the way that humans are. The traits selected by natural selection are described by Darwin as follows: “Man can act only on external and visible characters; nature cares nothing for appearances, except in so far as they may be useful to any being. She can act on every internal organ, on every shade of constitutional difference, on the whole machinery of life.” [35] so the process of natural selection favors any variant that increases the rate of survival or reproduction, even if the change is not easily perceptible by a human observer. Darwin thought of natural selection as gradual, but powerful. Over evolutionary time (generations), modifications would accumulate and cause substantially new forms to emerge. Natural selection could act even on small, imperceptible traits, if there was any difference in fitness. Moreover, Darwin conceived of natural selection as acting continuously, throughout all of nature, in a way that was for the most part imperceptible, befitting a process that could produce all biological life forms. It seems though that Darwin under-estimated the rate of natural selection, as we can see under certain circumstances (quickly-reproducing organisms like bacteria), its effects are detectable in a few years or less.

But for natural selection to operate, there must be many individuals that do not survive to reproduce. An early contribution of economic theory to evolution is a mechanism by which this happens: due to Malthus, it was shown that when there are finite resources, only a finite population size can be maintained. And so if one assumes the same for biological organisms, one obtains that there exist many of them which do not reproduce themselves. It happens that in the work presented in this thesis we also give contributions of economic theory to evolution.

Pre-Darwinian theories of evolution (like those of Lamarck) were based on a *transformational process*, by which the population changes due to the individuals’ coordinated individual changes. An example would be a potluck, by which each individual brings their dish, and the group of individuals now has an entire feast. By contrast, Darwin’s theory is based on a *variational process*, by which different variants are sorted out at a population level [80]. For example, to illustrate a variational process, consider a buffet, in which there are many different foods, and then the guests select their preferred dishes

from the mix. The remarkable insight that made this possible was a commitment by Darwin to an ever-changing world, bringing out things “ever ancient, ever new.” [8] Ernst Mayr [92] points out that in a particularly popular form of Platonism during the time of Darwin (which inspired some early approaches to taxonomy), there was an idea that all species must be fixed a priori. Though this static form of Platonic idealism is not representative of all Platonically-inspired metaphysics—especially in its Aristotelian form (see Oderberg’s work on the species concept in Aristotle [106])—clearly Darwin was advancing a theory directly opposed to it. A theory based on continual change.

Principles of Natural Selection

Having covered the Darwinian account of natural selection and all of its relevant assumptions, the contemporary account may be described. Wherever the following are present, then natural selection operates.

1. **Variation.** Individuals in a population differ in some ways from one another.
2. **Inheritance.** These differences are inheritable by descendants (offspring) from their ancestors (parents).
3. **Differential reproductive success.** Individuals with certain traits are more successful than others at surviving and reproducing in their habitat.

How can natural selection operate with just these three conditions? Let us keep in mind the following four points before we proceed in giving an example. First, an infrequent but continuous source of variation is mutations, which happen by chance at a small rate. There are other sources of variation that we will describe later as well. Second, evolutionary biologists study how perceptible traits evolve over time in order to find evidence of natural selection. Third, natural selection does not act on genotypes, but rather on phenotypic differences in the population. Therefore, to understand natural selection, we must understand how a phenotype is influenced by the genotype, the environment, and other factors. Fourth, natural selection directly changes the characteristics of a population, not of an individual.

Let us consider the example of natural selection acting on coat color for *Peromyscus polionotus*. The breed of mouse is commonly found in the Southeast, and preyed upon by visually hunting predators, including owls. Without variation, there is nothing for natural selection to act on. For coat color, in natural populations, there is observed variability within populations [99] in general, and at the *Mc1R* (melanocortin-1 receptor) locus associated with coat color. Even more variation is present between populations. The variation in phenotype could be due to variations in environment or genotype. But natural selection can operate only when genetic variability is present. We now explore why this is the case.

In *P. polionotus*, selection requires inheritance if it is to alter coat colors' frequencies in the next generation. Imagine for instance that light-colored mice produce five offspring on average, and dark-colored mice produce twelve offspring on average. If the offspring don't resemble their parents in coat color, then the dark-colored parents are just as likely to produce dark-colored offspring as light-colored parents. Thus the effects of differential reproductive success between the two groups will be eliminated after the parents produce new offspring. The mechanism for inheritance of a phenotypic variant is commonly the inheriting of a variant gene, that is correlated with the phenotypic change. The *Mc1R* gene is part of a genetic switch that controls the type of pigment created and incorporated into the hair. Another gene involved in coat color is called *Agouti*. The gene functions by having its product bind to the McR1 transmembrane receptor (which synthesizes darker pigment) and inhibiting its operation, thereby allowing for the increase in yellow pigment.

But genetic variation alone is not sufficient for natural selection to operate. Indeed, the variation must also have fitness consequences. Fitness has the quotidian implication of something that is well-suited—*fit*—to its world, but in evolutionary biology it connotes reproductive success. The fitness of a trait or allele is defined as the average reproductive success—of an individual who has that allele—relative to others in the population [47]. In the case of *P. Polionotus*' coat color, it impacts the fitness as follows: by increasing or decreasing visibility against their background. Then for visual predators, more visible mice are less difficult to find and thus more easily hunted.

With both of these conditions combined, the coat color thus can increase or decrease predation, and thus to survive long enough to reproduce.

The amazing fit between organisms and their environments motivated many of Darwin's predecessors to advance early evolutionary theories. An adaptation is an inherited trait that makes an organism more fit in its environment—whether abiotic (nonliving) or biotic (living). Note that something like a horse-shoe, though beneficial to a horse, is not an adaptation because it is not heritable. To be considered an adaptation, a trait must be shaped by natural selection to achieve the same primary function or functions that it has today. For example, the traits which were used for very different functions in the past as compared to the future, are called an *exaptation*.

It is important to remember that Evolution does not work in the way that an artificial selector would: it is a tinkerer, not a designer. Any change introduced will be very small in most cases. Imagine if you will how difficult it would be to generate a new operating system from a previous version, with random small changes introduced and evaluated by a committee for inclusion based on quality metrics. In a way this is how natural selection works: one can only identify good innovations in retrospect, based on their fitness. In this sense natural selection lacks foresight: it cannot choose the next step in any purposeful way. Indeed it is more the mode of variation that is most common in the population and, as Andreas Wagner points out, the space of possible phenotypes [140], that regulates the direction of selection. We will see that MWUA also shares this attribute when it selects between experts on the basis of the quality of their advice. As shown in the thesis, this resemblance is much more than just informal.

Given the amazing intricate structures that natural selection produces, how could such an accident-prone mechanism produce such great results? In other words, what is the mechanism by which natural selection could produce complex structures? For example, how could we explain the origins of the eye, a structure complex in many idiosyncratic ways? These questions are important for the thesis in that one of the papers covers the origin of tool-innovation, which truly requires novelty and the acquisition of incredibly complex skills. One possibility is that there are intermediates that

are partially-functional which tend towards the present-day form. The second possibility is that of exaptations, things that are complex but which served a different purpose before than they do now, in the present-day environment. For the intermediate theory of complexity, Ernst Mayr fleshed out a proposal by Darwin that the eye evolved from a layer of photoreceptor cells in molluscs to having a lens, thus explaining both the complexity of the structure and the presence of a blind-spot. For the exaptation theory, an example is the evolution of feathers. It is thought that feathers originally evolved for a variety of reasons, including heat retention, shielding from sunlight, signaling, prey capture, defense, and waterproofing [115]. In the function of retaining heat, feathers achieve this by acting as an insulator (feather down) and because the air space between feathers protects animals against temperature change. Especially exaptation is relevant to the origin of tool-innovation in the New Caledonian crow, as we will describe in Chapter 6.

2.1.2 Variation and Population Thinking

This subsection discusses the basics of genetics, then variation, and finally, population thinking in evolutionary biology. As mentioned earlier in the chapter, the basic unit of heredity is DNA, composed of base pairs. The genotype for an organism is the strip of DNA that it inherits from its parent(s). For natural selection to operate properly according to the previous section, we need heredity (DNA) but also the ability for the DNA to affect the fitness of the organism. For the fitness to be affected, it must be the case that the DNA somehow changes the phenotype. Otherwise, it would not have any impact on its powers/abilities as an organism [97], and thus would not affect the way the organism lives in the world—to its advantage or disadvantage. The process of translation from DNA to phenotypic change starts with transcription from functional DNA (introns) to RNA. RNA then directs protein synthesis (through mRNA in particular), by taking amino acids and putting them together in a characteristic sequence. That all of this is possible is due to a kind of code, that allows for the coding of various functions, that proteins have, by sequences of amino acids, and ultimately by codons (triplets of base pairs). The property of the genetic code that makes it so versatile

is its resilience to small errors through redundancy of amino acids' correspondence to proteins. But most of all, the genetic code's reliance on DNA makes it possible to have both combinatorial coding power (through the sequence of base pairs) and also the ability to be easily copied (through the hydrogen bonds in the center and the helical shape, allowing an "unzipping" for the transcription process).

One current definition of a gene is a sequence of DNA that specifies a functional product. Different variants of the same gene are called alleles, with the physical location(s) of a gene being called a locus(loci). A more rigorous definition of a genotype is the combination of alleles that an organism has at a given locus, or more generally at all loci. Organisms have either one copy of an allele at a locus, in which case they are called **haploid**, or they have two copies of every allele, in which case they are **diploid** organisms. Most organisms are diploid, but oftentimes for simplicity in mathematical models of evolution one assumes an organism is diploid. If diploid individuals have two copies of the same allele at the same locus, they are called homozygotes, and those with different alleles at a locus are called heterozygotes. Fr. Gregor Mendel, O.S.A., discovered the basics of genetics, introducing the surprising fact that changes in organisms were discrete and not continuous (contra Darwin). Like Darwin, he was interested in both the practical aspects of breeding, and also to some extent the metaphysics of species. In fact, Fr. Mendel viewed his work as investigation of the contribution of hybridization to the origins of novel organic forms [59], and his view of characters (phenotypes) was that they emerged in a radically top-down way similar to what is envisioned by Aristotle in his distinction between *potentia* and *actus* [70]. Briefly, whatever is potential in a given hybrid is whatever could be (and continues to be possible), whereas *actus* is the manifested and measurable. In more detail, St. Thomas Aquinas says "something is found to be in potency... in the sense that the whole can be reduced to act, as it is possible for this bronze to be a statue, because at some time it will be a statue" [108]. Moreover, Mendel's notion of dominant vs. recessive characters, and the fact that recessive characters could still be manifested in future generations means that recessive traits are potential manifestations [70].

As discussed in the last subsection, natural selection requires genetic variation. New

genetic variation enters a population from many possible sources, including recombination and mutation. First, we will describe recombination, then briefly explore mutation.

The standard way to introduce recombination is for diploid organisms, which (for eukaryotic organisms) have a fixed number of chromosomes. The chromosomes usually come in homologous pairs— pairs whose loci are the same, despite having possibly different alleles. One copy of each homologous pair of chromosomes comes from each parent— each contributes a gamete, a haploid cell with one set of chromosomes— in a process called meiosis. The gametes for the mother in the animal kingdom are eggs, and are sperm for the father. Mendel's first law, the law of segregation, states that each individual has two copies of a gene at each locus and that these gene copies segregate during gamete production—with only one gene copy making it into each gamete.

A dominant trait is one that if it appears in one copy of the gene at the locus or the other (inclusive or). A recessive trait appears only when both gene copies at a locus have that trait. Note then that each gene copy retains therefore its individual character, whether or not it manifests in any phenotypic change.

Meiosis starts out with one diploid cell; a single round of DNA replication is followed by two rounds of division which then gives four haploid gametes. When fertilization occurs (the two gametes from the parents fuse after mating), the resulting product is again diploid. The offspring will have all the necessary chromosomes, but with one chromosome in each pair coming from each parent. There are two kinds of reproduction we will discuss in the thesis: asexual reproduction and sexual reproduction. The former is what cells do: make a clone, and the latter requires two parents and involves exchanging genetic material during mating.

What does this have to do with variation? Sexually-reproducing organisms end up producing high genetic variability during meiosis, by the process of **cross-over**— the exchange of segments of DNA. Cross-over is a kind of recombination. During cross-over, areas of homologous chromosomes are exchanged as follows:

1. The homologous chromosomes are duplicated
2. Sections of one homologous chromosome may swap positions with sections on the

other homologous chromosome during meiosis. This happens at some fixed rate.

3. The resulting product is four daughter cells, consisting of chromosomes that could differ from those in the parents.

For haploid organisms, recombination is similarly a kind of shuffling of the genetic material of the parents. To scale up the effect of recombination to its effect on the whole population, consider that one should specify the pattern of mating in order to determine the final make-up of the population. A common assumption is that of panmixis— random mating pairs taken from the whole population.

Recombination generates new combinations based on the genotypes of parents, but how can one get a new set of parent genotypes? Mutation is the source of novel genetic variation. It is the change in the DNA sequence. One can classify the different kinds of mutations based on their effects on amino acids, for example, missense mutations cause the production of novel amino acids after a single change in base pair (a so-called point mutation). For other kinds of mutations, one can also add a base-pair to the DNA sequence (insertion mutation), and remove a base-pair (deletion mutation). There are a variety of other kinds of mutations, but this is sufficient to give a crucial intuition about mutation: it is a very small change to the DNA. But such a small change could have profound effects in the case of a missense mutation for instance, if the amino acid changed is crucial for proper functioning of the organism. An important factor to consider is the relative scarcity of mutations, they typically happen at a rate of 10^{-6} mutations per generation. The scarcity of mutations makes sources of variation like recombination more important, because they happen much more often. Luria and Delbruck [88] showed with an ingenious experiment that mutation happened in a way that cannot reliably be predicted on the basis of natural selection. The randomness accorded to the mutation process is important for the generation of new variants.

As mentioned earlier, an essential part of how evolutionary theory works now depends on thinking about a population of organisms changing, rather than individuals. The frequency of a given allele or genotype gradually changes over time, as the forces of variation and selection operate on the population. When a genotype dominates the

population, it is said to have fixated. If there are multiple genotypes in the population that are stable under the action of variation and natural selection, this is called polymorphism. As is clear upon a little reflection, having a finite population leads to sampling effects—commonly called genetic drift—assuming that one samples the same number of individuals from the current population’s descendants to choose who will be in the next generation (with each individual reproducing at a different rate, if left unchecked, the population size will grow without bound). One could for instance sample individuals that are by and large of low fitness, and thereby cause fixation of a deleterious mutant.

2.2 Population Genetics

Best of all in evolutionary biology, there are mathematical models to make rigorous all of the models described in the last section. First we will describe the simplest model: a haploid organism which reproduces asexually, with a single locus and no mutation. We will build up from there to various more complex models, culminating in models with two loci and recombination.

Consider a single locus with alleles A_i , where $i = 1, 2, \dots, k$, and assume that the generations happen in discrete time (discrete) and do not overlap (non-overlapping). What this amounts to is assuming the adults are replaced by their offspring in one generation. Let the number of offspring that have allele A_i in generation t , be $n_i(t)$. The total number of offspring:

$$N = \sum_i n_i$$

must be sufficiently large that we can make the “infinite population size” assumption. Such an assumption assures that the sampling effects of drift are negligible, and that we can instead just use the expected number of progeny, rather than the actual, random number. Let v_i be the probability that an A_i offspring survives to reproduce. The average number of descendants of an A_i adult is f_i . These are called viabilities v_i and fertilities f_i , which could be a function of time or population numbers n_i . The product $w_i = v_i f_i$ is the fitness of an A_i -carrier. There are $v_i n_i$ adults of type A_i , which

contribute $w_i n_i$ offspring to the next generation. Therefore, we have:

$$n_i(t+1) = w_i(t, n) n_i(t)$$

If the allele A_i is lethal, then $v_i = 0$, and if it causes sterility then $f_i = 0$. Otherwise, the w_i will not be too different, and on average will be close to unity for constant population size. A particular case of this we call weak selection: $w_i = (1 + s)\Delta_i$, with $s \rightarrow 0$ and Δ_i is the differential fitness. Let the selection coefficients s_i be $w_i - 1$. The particular values of the selection coefficients s_i for all i are called collectively the “fitness landscape”. Occasionally, the values of the fitnesses w_i themselves are called the “fitness landscape.” We will have both of these usages in the thesis.

The frequency of the allele A_i is

$$p_i = \frac{n_i}{N}$$

We obtain from the last three equations the recurrence relation:

$$N' = \sum_i n'_i = \sum_i w_i \left(\frac{n_i}{N}\right) N = \bar{w}N$$

where $\bar{w} = \sum_i w_i p_i$ is the average or mean fitness of the population. The frequencies satisfy the recurrence:

$$p'_i = \frac{n'_i}{N'} = \frac{w_i n_i}{\bar{w} N} = p_i \frac{w_i}{\bar{w}}$$

Clearly, $\sum_i p_i = 1$ for all generations, and gene frequencies only depend on the ratios of the fitnesses. Moreover, the dynamics are unchanged by a multiplicative scaling of the fitnesses. The difference equation form of the recurrence is given by:

$$\Delta p_i = p'_i - p_i = \frac{p_i(w_i - \bar{w})}{\bar{w}}$$

If the fitnesses are only dependent on time, then:

$$n_i(t) = n_i(0) \prod_{\tau=0}^{t-1} w_i(\tau)$$

Then the frequencies can be written as

$$p_i(t) = \frac{p_i(0)\prod_{\tau=0}^{t-1}w_i(\tau)}{\sum_j p_j(0)\prod_{\tau=0}^{t-1}w_j(\tau)}$$

Assuming that the time-scale at which the fitnesses vary are sufficiently slow relative to the other relevant evolutionary parameters, we can assume that the fitness is constant over time, obtaining:

$$n_i(t) = n_i(0)w_i^t$$

$$p_i(t) = \frac{p_i(0)w_i^t}{\sum_j p_j(0)w_j^t}$$

Let $w_1 > w_i$ for all $i > 1$. So A_1 is the fittest allele. Then $p_1(t) \rightarrow 1$ as $t \rightarrow \infty$. Then if the population size remains finite, fixation will happen for A_1 and the other alleles are lost. But of course all the preceding holds as long as $p_1(0) > 0$, when the allele A_1 is initially present in the population. Now we will see how in fact the average fitness \bar{w} actually increases over time under some conditions, a result called ‘‘Fisher’s fundamental theorem of Natural Selection,’’ after Ronald Fisher, a father of population genetics. For the change in average fitness, we have:

$$\begin{aligned} \Delta\bar{w} &= \sum_i (p'_i w'_i - p_i w_i) \\ &= \sum_i [p'_i (\Delta w_i + w_i) - p_i w_i] \\ &= \Delta\bar{w} + \sum_i w_i \Delta p_i \end{aligned}$$

where $\Delta\bar{w} = \sum_i p'_i \Delta w_i$ is the average change in fitness. Via a simple substitution of the analytic Δp_i into the last expression one obtains:

$$\Delta\bar{w} = \Delta\bar{w} + \bar{w}^{-1} \sum_i p_i w_i (w_i - \bar{w})$$

But by the definition of \bar{w} it holds that

$$\sum_i p_i (w_i - \bar{w}) = 0$$

So we can subtract \bar{w} from the first w_i in the definition of \bar{w} , and then get:

$$\Delta\bar{w} = \bar{\Delta}w + \bar{w}^{-1}V$$

where $V = \sum_i p_i (w_i - \bar{w})^2$ is the genic variance in fitness. With constant fitness, $\bar{\Delta}w = 0$, therefore $\Delta\bar{w} = \bar{w}^{-1}V \geq 0$, and thus the average fitness is nondecreasing. Therefore $\Delta\bar{w} = 0$ only at equilibrium, since $V = 0$ if and only if $p_i = 0$ or $w_i = \bar{w}$ for all i . Therefore natural selection ends up increasing the average fitness, at a rate determined by the genic variance. Because it is implausible to suppose that absolute fitness (actual fitness) will remain constant, it is more profitable to scale the fitnesses to be relative fitnesses, leaving the dynamics invariant, but then allowing the more plausible assumption that relative fitnesses remain constant. Doing so we can derive a similar uptick in average relative fitness given the preceding. To see this, note that we can take $w_i = z_i g(N)$, with z_i constant and $w_i < 1$ for sufficiently large N for all i .

Now let's explore the significance of the finding to biology. Let A_1 be the fittest genotype as before. Then natural selection minimizes the quantity $(w_1 - \bar{w})/w_1 = (z_1 - \bar{z})/z_1$, the relative reproductive excess of the optimal genotype. Note that the relative reproductive excess is also called the genetic load by Crow [31]. If the s_i are small, then we may choose the relative fitnesses to be close to unity. Then the average fitness will be close to unity, and its rate of change will be proportionate to the genic variance V . As it turns out the genetic load is a kind of measure of how well evolution increases the fitness of the population as compared to the best possible allele (in retrospect). In decision theory (assuming the fitness is the utility), such a quantity is called regret. We show that there are algorithms that, in a general-purpose way, minimize regret. One of them is MWUA.

Now we will discuss the case of two loci, first with just recombination and panmixis for diploid organisms, then later with selection. Let the alleles at A - and B -loci,

A_i and B_i , have frequencies p_i and q_i . Denote the frequency of A_iB_j gametes in the gametic output of generation t $P_{ij}(t)$. The gene frequencies are:

$$p_i = \sum_j P_{ij} \quad (2.1)$$

$$q_j = \sum_i P_{ij} \quad (2.2)$$

but the joint frequencies P_{ij} are no longer properly described by the gene frequencies. Gametes A_iB_j and A_kB_l form the individual A_iB_j/A_kB_l . A proportion $1 - r$ of the gametes produced by this individual are nonrecombinant (or parental) gametes, $\frac{1}{2}(1 - r)$ are of type A_iB_j and $\frac{1}{2}(1 - r)$ are of type A_kB_l , a fraction r are recombinant (or parental), i.e., $\frac{1}{2}r$ of type A_iB_l and the same number of type A_kB_j . When the two loci are on the same chromosome, $r \in [0, 1/2]$, where r is the recombination fraction. The value of r depends on the position of the two loci and the chromosomal structure. Assume that $r > 0$ (as we are not interested here in the single locus case, when $r = 0$). For loci on different chromosomes, $r = 1/2$. Unlinked loci have $r = 1/2$.

The proportion $1 - r$ of the gametes in generation $t + 1$ are produced without recombination. Thus the expected contribution to the new joint frequencies is $(1 - r)P_{ij}$. Now with random mating, the expected contribution to the new joint frequencies of the recombinant events is rp_iq_j . Therefore, by linearity of expectation:

$$P'_{ij} = (1 - r)P_{ij} + rp_iq_j \quad (2.3)$$

It follows from Equations 2.1 and 2.3 that the gene frequencies are still normalized. Now let the linkage disequilibria be:

$$D_{ij} = P_{ij} - p_iq_j$$

which is a measure of the departure from random combination of alleles within gametes.

From the definition of linkage disequilibrium combined with Equation 2.3, we obtain:

$$D'_{ij} = (1 - r)D_{ij} \quad (2.4)$$

and therefore

$$D_{ij}(t) = D_{ij}(0)(1 - r)^t$$

so $D_{ij} \rightarrow 0$. The linkage equilibrium $P_{ij} = p_i q_j$ is attained gradually, and faster for larger values of r . The ordered frequency of genotype $A_i B_j / A_k B_l$ is $P_{ij} P_{kl}$. By summing $P_{ij} P_{kl}$ over j and l , or i and k , we derive at each locus the quantity called the Hardy-Weinberg equilibrium. At linkage equilibrium, $P_{ij} P_{kl}$ is $(p_i p_k)(q_j q_l)$, the product of the single-locus Hardy-Weinberg frequencies.

Now let us describe the same model with selection, using similar notation. The fitnesses of $A_i B_j / A_k B_l$ and $A_i B_l / A_k B_j$ are often assumed to be the same, so $w_{ij,kl} = w_{il,kj}$ (which is called a lack of position effect). But we need not assume this. Let $R_{ij:kl,mn}$ be the probability that $A_i B_j$ is a gamete produced at random by an $A_i B_j / A_k B_l$ individual. Then

$$P'_{ij} = \bar{w}^{-1} \sum_{klmn} w_{kl,mn} P_{kl} P_{mn} R_{ij:kl,mn} \quad (2.5)$$

with

$$\bar{w} = \sum_{klmn} w_{kl,mn} P_{kl} P_{mn} \quad (2.6)$$

Let r be the recombination fraction between A and B loci. To calculate R , draw A_i , and take into account recombination:

$$R_{ij:kl,mn} = \frac{1}{2} \delta_{ik} [(1 - r)\delta_{jl} + r\delta_{jn}] + \frac{1}{2} \delta_{im} [(1 - r)\delta_{jn} + r\delta_{jl}] \quad (2.7)$$

The sum in Equation 2.5 is invariant to swaps between k and m , and l and n . Therefore the two terms in Equation 2.7 contribute equally. Using this information, we derive:

$$\bar{w} P'_{ij} = P_{ij} w_{ij} - r D_{ij} \quad (2.8)$$

where the linkage disequilibria are

$$\bar{w}D_{ij} = \sum_{kl} (w_{ij,kl}P_{ij}P_{kl} - w_{il,kj}P_{il}P_{kj}) \quad (2.9)$$

and

$$w_{ij} = \sum_k l w_{ij,kl} P_{kl}$$

is the fitness of a gamete $A_i B_j$. The linkage disequilibria satisfy

$$\sum_i D_{ij} = 0$$

and

$$\sum_j D_{ij} = 0$$

Now we define linkage equilibrium in a similar way as before, the random combination of alleles within gametes: $P_{ij} = p_i q_j$ for all i, j , where

$$p_i = \sum_j P_{ij} \quad (2.10)$$

$$q_j = \sum_i P_{ij} \quad (2.11)$$

are the frequencies of A_i and B_j respectively. Given linkage equilibrium, if there are no position effects, Equation 2.9 yields $D_{ij} = 0$ for all i and j . Furthermore, in most of the use of this model, we will assume $s \ll r$. Finally, the haploid model is a special case of this one, in which $w_{ij,kl} = w_{ij} w_{kl}$.

2.3 Mixability

A big motivator for mixability theory was the frequent observation that genetic algorithms are not always effective at optimizing functions. The natural question to ask was, what if genetic algorithms (with recombination) are not optimizing a straightforward fitness function?

In fact, recombination can sometimes break-up good combinations of alleles, and so

it can actually move one away from a fitness peak!

What does it maximize then? In the short run, the so-called “transient,” it increases the frequency of alleles that mix well with a free choice of the allele for the other locus. In other words, it increases the average marginal fitness of an allele, assuming equal frequency (uniform distribution) of the alleles in the population.

Livnat and Papadimitriou show using simulations that the advantage to mixability is only transient, and eventually goes away. This suggests a role for mathematics to analyze the transient properties of dynamical systems, as typically the Population Genetics community focuses on the equilibria or long-term behavior.

What kind of thing does mixability measure? It is like hedging your bets on the stock market. Mixability is choosing an allele that does very well in case there is an unexpected genetic change.

Mixability was defined in the context of the mystery as to why sexual reproduction evolved at all. There is a well-known flaw in sexual reproduction relative to asexual reproduction: that in fact it requires more individuals to produce the same number of offspring, and could in fact break apart good combinations of alleles, thereby reducing fitness. So therefore, it isn't just of interest to those pondering why and how genetic algorithms fail— but may be fundamental to understanding genetics itself.

Chapter 3

The Multiplicative Weight Updates method

In this chapter, we describe background information on the Multiplicative Weight Updates Algorithm (MWUA) and also its application to solving computational problems (the MWU method). Then we will apply the MWU method to solving a zero-sum game.

3.1 Multiplicative Weight Updates

Consider a most general scenario: we have a set of n decisions and on each round, we must select one decision from the set. Each decision incurs a cost, determined by nature or an adversary. Costs are all revealed after decisions are chosen, and the cost of the decision is incurred at that time. A simple enough strategy is to, in each iteration, pick a decision at random. The incurred cost will be that of the “average” decision. Suppose that there are a few decisions that are much better in the long run. Assuming that nature is relatively stable, then it is easy to see after a short time which decisions are best based on costs over time. Therefore it is sensible to reward the good decisions by increasing their probability of being picked in the next round.

Being ignorant about the decisions at the outset, they are selected initially at random. The entropy is maximized for this decision rule, as we will prove later. Such an initial rule signals our lack of information. As it becomes more apparent which decisions are good and which ones are bad, we lower the entropy to reflect the increase in information. It is important to note that there is embedded in this theoretical framework an assumption that each decision’s outcome actually gives us knowledge about which one is best in the long run.

Now let us setup a more formal discussion, according to the MWUA as presented by

Arora et al. [6]. There are n experts providing advice at each of T time steps. Listening to expert $i \in n$ at time step $t \in T$ will yield the loss $l_i^{(t)} \in [-1, 1]$. For example, one can think of these experts as meteorologists predicting the weather, and the loss is their error. The goal is to find a strategy for paying attention to experts such that, after many time steps, the strategy will perform almost as well as a strategy that listened to the expert which gave the best advice over the whole time period, without knowing in advance the losses of the experts (no assumption is made about the losses, they may be correlated, or even chosen adversarially).

The expected loss to the algorithm for sampling an expert's decision i from the distribution $p^{(t)}$ is

$$E_{i \in p^{(t)}}[l_i^{(t)}] = l^{(t)} \cdot p^{(t)}$$

And thus the total expected loss over all rounds is $\sum_{t=1}^T l^{(t)} \cdot p^{(t)}$. Our goal is to make an algorithm which performs well in total expected loss compared to the best decision in hindsight, viz. $\min_i \sum_{t=1}^T l_i^{(t)}$.

Such a strategy exists, and the MWUA is it. The MWUA assigns a weight w_i to each expert i . At each time step, an expert is chosen with a probability proportional to its weight. The vector of advice-following probabilities at time step t is then $p^{(t)} = \{w_1^{(t)}/\Phi^{(t)}, \dots, w_n^{(t)}/\Phi^{(t)}\}$, where $\Phi^{(t)} = \sum_i w_i^{(t)}$. The losses of the different experts are then observed, and the weights are updated according to the following rule: $w_i^{(t+1)} = w_i^{(t)}(1 - \eta l_i^{(t)})$, using a fixed $\eta > 0$.

The disparity between the total expected loss and that of the best expert in hindsight is called “regret.” Which we can quantify as:

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} - \min_i \left(\sum_{t=1}^T l_i^{(t)} + \eta \sum_{t=1}^T |l_i^{(t)}| \right)$$

Theorem 2.1 in Arora et al. states that, under the MWUA (when $l_i^{(t)} \in [-1, 1]$ and $\eta \leq 1/2$), for any agent i :

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} \leq \sum_{t=1}^T l_i^{(t)} - \eta \sum_{t=1}^T |l_i^{(t)}| - \frac{\ln n}{\eta}.$$

For large T , the last term becomes negligible, and for small η , the second term is small. Now, since this inequality holds for any expert i , it holds specifically also for the best expert in retrospect, and thus the performance of the MWUA is close to that of the best expert in retrospect.

Now we will prove a no-regret version of Kale's theorem.... The key intuition for how it works is that one uses a potential function to track progress, and track how it changes over a single step. Then use the resulting upper- and lower-bounds to relate the total expected loss and the minimum loss. What makes this work ultimately is the exponential function converting addition of its arguments into multiplication of the exponential of each argument separately.

Theorem 1. *Using MWUA with $l_i^{(t)} \in [-1, 1]$ and $\eta \leq 1/2$, for any agent i :*

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} \leq \min_i \left(\sum_{t=1}^T l_i^{(t)} - \eta \sum_{t=1}^T |l_i^{(t)}| \right) - \frac{\ln n}{\eta}$$

Proof. The proof uses the potential function $\Phi_t = \sum_i w_i^{(t)}$, and the fact that $p_i^{(t)} = w_i^{(t)} / \Phi^{(t)}$:

$$\begin{aligned} \Phi_{t+1} &= \sum_i w_i^{(t+1)} \\ &= \sum_i w_i^{(t)} (1 - \eta l_i^{(t)}) \\ &= \Phi_t - \eta \Phi_t \sum_i l_i^{(t)} p_i^{(t)} \\ &= \Phi_t (1 - \eta l^{(t)} \cdot p^{(t)}) \\ &\leq \Phi_t \exp(-\eta l^{(t)} \cdot p^{(t)}) \end{aligned}$$

Then by induction, after T rounds, we have

$$\Phi_{T+1} \leq \Phi_1 \exp(-\eta \sum_{t=1}^T l^{(t)} \cdot p^{(t)}) = n \exp(-\eta \sum_{t=1}^T l^{(t)} \cdot p^{(t)}) \quad (3.1)$$

Next it is necessary to upper-bound the exponentials with linear functions as follows:

$$(1 - \eta)^x \leq (1 - \eta x) \quad (3.2)$$

if $x \in [0, 1]$, and

$$(1 + \eta)^{-x} \leq (1 - \eta x) \quad (3.3)$$

if $x \in [-1, 0]$. Since $l_i \in [-1, 1]$, then, for every expert i ,

$$\Phi_{T+1} \geq w_i^{(T+1)} = \prod_{t \leq T} (1 - \eta l_i^{(t)}) \geq (1 - \eta)^{\sum_{\geq 0} l_i^{(t)}} \cdot (1 + \eta)^{\sum_{< 0} l_i^{(t)}} \quad (3.4)$$

where the subscripts “ ≥ 0 ” and “ < 0 ” in the summations refer to the rounds t where $l_i^{(t)}$ is ≥ 0 and < 0 respectively. By taking logarithms in equations 3.4 and 3.1, it follows that:

$$\ln n - \eta \sum_{t=1}^T l^{(t)} \cdot p^{(t)} \geq \sum_{\geq 0} l_i^{(t)} \ln(1 - \eta) - \sum_{< 0} l_i^{(t)} \ln(1 + \eta)$$

Negating, rearranging, and scaling by $1/\eta$, we get:

$$\begin{aligned} \sum_{t=1}^T l^{(t)} \cdot p^{(t)} &\leq \frac{\ln n}{\eta} + \frac{1}{\eta} \sum_{/ge 0} l_i^{(t)} \ln \frac{1}{1 - \eta} + \frac{1}{\eta} \sum_{< 0} l_i^{(t)} \ln(1 + \eta) \\ &\leq \frac{\ln n}{\eta} + \frac{1}{\eta} \sum_{\geq 0} l_i^{(t)} (\eta + \eta^2) + \frac{1}{\eta} \sum_{< 0} l_i^{(t)} (\eta - \eta^2) \\ &= \frac{\ln n}{\eta} + \sum_{t=1}^T l_i^{(t)} + \eta \sum_{t=1}^T |l_i^{(t)}| \end{aligned}$$

To derive the second inequality the following two inequalities were used:

$$\begin{aligned} \ln \left(\frac{1}{1 - \eta} \right) &\leq \eta + \eta^2 \\ \ln(1 + \eta) &\leq \eta - \eta^2 \end{aligned}$$

for $\eta \leq 1/2$. Then since the inequality holds for all i , we can choose the i which minimizes the total loss, whence the result. \square

Notably, though there are other variants of the MWUA using the exponential form

$\exp(-\eta x)$, this variant has the strongest guarantees: the term multiplying η is a loss which depends on the loss of the best expert in hindsight. Which is crucial for applications of the algorithm to various domains [6]

There is also a version of MWUA for gains instead of losses:

The Multiplicative Weight Update Algorithm (MWUA) [6] chooses between k different experts, choosing from them according to some probability distribution according to how poor their predictions are. The quality of the prediction for an expert i at time t is its gain $g_i^{(t)}$. The probability distribution over experts $p_i^{(t)}$ at time $t + 1$ for MWUA is given by:

$$p_i^{(t+1)} = p_i^{(t)} \frac{1 + \eta g_i^{(t)}}{\sum_j p_j^{(t)} (1 + \eta g_j^{(t)})}$$

with $\eta > 0$ being the learning rate. In words, the experts with worse predictions are chosen with a lower probability at time $t + 1$ than at time t , and this is done more aggressively when η is larger.

And from this follows:

Corollary 2. *The Multiplicative Weights algorithm also guarantees that after T rounds, for any distribution p on the decisions:*

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} \leq \sum_{t=1}^T (l^{(t)} + \eta |l^{(t)}|) \cdot p + \ln n / \eta$$

where $|l^{(t)}|$ is the vector obtained by the taking the coordinate-wise absolute value of $l^{(t)}$.

There is a corresponding result that bounds the regret for the gains version. One obtains it by using an identical argument as for the losses case, and plugging in $l^{(t)} = -g^{(t)}$ into all of the bounds. One obtains:

Theorem 3. *Using MWUA with $g_i^{(t)} \in [-1, 1]$ and $\eta \leq 1$, for any agent i , we obtain the following:*

$$\sum_{t=1}^T g^{(t)} \cdot p^{(t)} \geq \max_i \left(\sum_{t=1}^T g_i^{(t)} - \eta \sum_{t=1}^T |g_i^{(t)}| \right) - \frac{\ln n}{\eta}$$

3.1.1 MWUA as a Game dynamics

In a game-theoretic setting, MWU can be used in a setting involving two algorithms whose losses depend on each other. But first let us review a little bit of game theory. In a 2-player game, there are players Row and Column. Each player plays a pure strategy: for row player, a row index i , for column player, a column index j . But also players could choose a mixed strategy: a probability distribution over pure strategies for the player. And then the payoffs are associated with a choice of row and column for each player, with both players having their own payoff matrix. A bimatrix game is $G = (A, B, [n], [m])$ where A is the payoff (utility) matrix of the row player, B the payoff (utility) matrix of the column player, n is the number of possible rows for the row player, and k is the number of possible columns for the column player. Here are two examples of bimatrix games:

Coordination games have identical payoff matrices for both players, so $A = B$ and $n = m$. Consider Generalized Rock-paper-scissors games. A is

$$\begin{pmatrix} 0 & -b & a \\ a & 0 & -b \\ -b & a & 0 \end{pmatrix}$$

and $B = A^T$. For zero-sum games $A = -B$, and $\dim(A) = \dim(B) = n$.

The application for MWUA in game theory is an iterative algorithm to get a pair of mixed strategies that do very well in terms of expected payoff (given the other player's mixed strategy). The general form for game dynamics in Bimatrix games is given by:

$$r_{t+1}^i = f(r_t, [Ac_t^T], \theta, i)$$

$$c_{t+1}^j = f(c_t, [r_t^T B], \theta, j)$$

In words, a game dynamics will take the vector of expected payoffs for different pure strategies and adjust the mixed strategy to increase the probability of choosing that strategy accordingly.

For MWUA to be applied as a game dynamics, one just takes

$$f(v, u, \eta, i) = v_i \frac{1 + \eta u_i}{\sum_j v_j (1 + \eta u_j)}$$

In Algorithms, MWUA is used as a general-purpose optimization algorithm, as a way to solve many different optimization problems. To use the method, one need only specify the loss functions and experts in such a way as to solve the problem. We will describe this method in the following subsection.

3.1.2 The MWU Method and its Applications

The MWU method is the application of the MWUA to a given function problem, supplying its inputs to the MWUA as experts and a corresponding set of losses for the experts, such that the output is a solution to the function problem. More formally, an application of the MWU method supplies a set of experts $i \in [n]$, and a set of loss vectors $l^{(t)}$ for each expert on each time-step. The output of MWUA on the provided set of experts/losses after a long-enough time will then need to be close-to optimal for the given problem.

For an example of an application of the MWU method, consider approximately-solving a 2-player zero-sum game.

Let A be the payoff matrix of a bimatrix 2-player zero-sum game. Recall that the expected payoff to the column player for choosing a strategy j is $A(p, j) = E_{i \in p}[A(i, j)]$. Therefore the best response for the column player is to choose the strategy j which maximizes this payoff. If the column player chooses his strategy j from a distribution over the columns is $A(i, q) = E_{j \in q}[A(i, j)]$. Thus the best response for the row player is the row which minimizes the expected payoff (recall $B = -A$). John von Neumann's min-max theorem says that if each of the players choose a mixed strategy to optimize their worst-case payoff, then the value they obtain is the same:

$$\min_p \max_j (A(p, j)) = \max_q \min_i A(i, q) \tag{3.5}$$

where p, q , vary over all mixed strategies for the row and column players respectively. Also, i (resp. j) varies over all rows (resp. columns). The common value of both sides of the equation, λ^* , is called the value of the game. To solve a game up to additive error $\epsilon > 0$, we must find mixed row and column strategies \hat{p} and \hat{q} such that:

$$\lambda^* - \epsilon \leq \min_i A(i, \hat{q}) \tag{3.6}$$

$$\max_j A(\hat{p}, j) \leq \lambda^* + \epsilon \tag{3.7}$$

Let us assume that given any distribution p on strategies, we have an efficient way to pick the best response, e.g. the pure strategy j that maximizes $A(p, j)$, which is at least λ^* . Call this algorithm Oracle.

Theorem 4. *Given an error parameter $\epsilon > 0$, one can apply the MWU method up to solve the zero-sum game up to an additive factor of ϵ using $O(\log(n)/\epsilon^2)$ time-steps (and calls to Oracle), with an additional processing time of $O(n)$ per call.*

Proof. We consider 3.6 to specify n linear constraints on the probability vector \hat{q} : viz., for all rows i , $A(i, \hat{q}) \geq \lambda^* - \epsilon$. Let the experts be the pure strategies of the row player. Similar to the MWUA in a game-theoretic setting. Losses are specified by pure strategies of the column player. The loss for expert i when column player chooses strategy j is $A(i, j)$.

In each time-step, given a distribution $p^{(t)}$ on the rows, we choose the column $j^{(t)}$ to be the best response strategy to $p^{(t)}$ for the column player using a call to Oracle. The loss for expert j is thus the j^{th} column of A . Similar to the MWU in a game-setting, assuming that the column player always plays the best response.

Since all $A(i, j) \in [0, 1]$ by re-normalization, we can apply Corollary 2 to get that after T time-steps, for any distribution on the rows p , we have

$$\sum_{t=1}^T A(p^{(t)}, j^{(t)}) \leq (1 + \eta) \sum_{t=1}^T A(p, j^{(t)}) + \ln n/\eta$$

Dividing by T , and using the fact that $A(p, j) \leq 1$ and that for all t , $A(p^{(t)}, j^{(t)}) \geq \lambda^*$, we obtain:

$$\lambda^* \leq \frac{1}{T} \sum_{t=1}^T A(p^{(t)}, j^{(t)}) \leq \frac{1}{T} \sum_{t=1}^T A(p, j^{(t)}) + \eta + \frac{\ln n}{\eta T}$$

Setting $p = p^*$, the optimal row strategy, we have $A(p, j) \leq \lambda^*$ for any j . By setting $\eta = \epsilon/2$ and $T = \lceil 4 \ln(n)/\epsilon^2 \rceil$, we get that

$$\lambda^* \leq \frac{1}{T} \sum_{t=1}^T A(p^{(t)}, j^{(t)}) \leq \sum_{t=1}^T A(p, j^{(t)}) \leq \lambda^* + \epsilon \quad (3.8)$$

Thus, $\frac{1}{T} \sum_{t=1}^T A(p^{(t)}, j^{(t)})$ is an additive ϵ -approximation to λ^* . Let \hat{t} be the round t with the minimum value of $A(p^{(t)}, j^{(t)})$. We have, from the above,

$$A(p^{(\hat{t})}, j^{(\hat{t})}) \leq \frac{1}{T} \sum_{t=1}^T A(p^{(t)}, j^{(t)}) \leq \lambda^* + \epsilon$$

Since $j^{(\hat{t})}$ maximizes $A(p^{(\hat{t})}, j)$ over all j , we conclude that $p^{(\hat{t})}$ is an approximately optimal mixed strategy for the row player, and thus we can set $p^* = p^{(\hat{t})}$.

We set q^* to be the distribution which assigns to column j the probability

$$\frac{|t : j^{(t)} = j|}{T}$$

From 3.8, for any row strategy i , by setting p to be concentrated on the pure strategy i , we have

$$\lambda^* - \epsilon \leq \frac{1}{T} \sum_{t=1}^T A(p^{(t)}, j^{(t)}) = A(i, q^*)$$

which shows that q^* is an approximately-optimal mixed strategy for the column player. \square

In fact, one can use a very similar analysis to show that the straightforward application of MWUA for games outlined in the previous subsection gives nearly the same guarantees. There are many other examples given by Kale and co-authors in his review of applications for the MWU method, including solving convex programming, linear

programming, classification problems, and some simple graph-theoretic problems [6].

Chapter 4

Population Genetics and MWUA

In this chapter, we cover joint work with Papadimitriou, Livnat, and Vazirani on an unexpected connection between population genetics and MWUA playing a coordination game. The writing in this chapter was done by myself, Papadimitriou, and Livnat. Technical and modeling contributions were given by all four authors. The primary content of this paper has been published elsewhere [25, 24], but the last section on mixability is so far unique to the thesis. Recall the definition of a game. A coordination game has payoff matrices A, B such that $B = A^T$. Now we will present the basic connection to MWUA in both genetic and game-theoretic terms. We assume that the genes (pure strategies) are the experts, and the average fitness for each allele (payoff to each pure strategy) is $-l_i$.

4.1 MWUA and Population Genetics

Precisely how does selection change the composition of the gene pool from generation to generation? The field of population genetics has developed a comprehensive mathematical framework for answering this and related questions [21]. Our analysis in this paper focuses particularly on the regime of weak selection, now a widely used assumption [100, 104]. Weak selection assumes that the differences in fitness between genotypes are small, and consequently, through a result due to Nagylaki et al. [103], (see also [21] Section II.6.2), evolution proceeds near linkage equilibrium, a regime where the probability of occurrence of a certain genotype involving various alleles is simply the product of the probabilities of each of its alleles. Based on this result, we show that evolution in the regime of weak selection can be formulated as a repeated game, where the recombining loci are the players, the alleles in those loci are the possible actions or

strategies available to each player, and the expected payoff at each generation is the expected fitness of an organism across the genotypes that are present in the population. Moreover, and perhaps most importantly, we show that the equations of population genetic dynamics are mathematically equivalent to positing that each locus selects a probability distribution on alleles according to a particular rule which, in the context of the theory of algorithms, game theory and machine learning, is known as multiplicative weight updates (MWUA). MWUA is known in computer science as a simple but surprisingly powerful algorithm (see [6] for a survey). Moreover, there is a dual view of this algorithm: each locus may be seen as selecting its new allele distribution at each generation so as to maximize a certain convex combination of (a) cumulative expected fitness and (b) the entropy of its distribution on alleles. This connection between evolution, game theory, and algorithms seems to us rife with productive insights; for example, the dual view just mentioned sheds new light on the maintenance of diversity in evolution.

Game theory has been applied to evolutionary theory before, to study the evolution of strategic individual behavior (see, e.g., [91, 144]). The connection between game theory and evolution that we point out here is at a different level, and arises not in the analysis of strategic individual behavior, but rather in the analysis of the basic population genetic dynamics in the presence of sexual reproduction. The main ingredients of evolutionary game theory, namely strategic individual behavior and conflict between individuals, are extraneous to our analysis.

We now state our assumptions and results. We consider an infinite panmictic population of haplotypes involving several unlinked (i.e., fully recombining) loci, where each locus has several alleles. These assumptions are rather standard in the literature. They are made here in order to simplify exposition and algebra, and there is no a priori reason to believe that they are essential for the results, beyond making them easily accessible. For example, Nagylaki's theorem [103], which is the main analytical ingredient of our results, holds even in the presence of diploidy and partial recombination.

Nagylaki's theorem states that weak selection in the presence of sex proceeds near

the Wright manifold, where the population genetic dynamics become (SI)

$$x_i^{t+1}(j) = \frac{1}{X^t} x_i^t(j) (F_i^t(j)),$$

where $x_i^t(j)$ is the frequency of allele j of locus i in the population at generation t , X is a normalizing constant to keep the frequencies summing to one, and $F_i^t(j)$ is the mean fitness at time t amongst genotypes that contain allele j at locus i (see [103] and the SI). Under the assumption of weak selection, the fitness of all genotypes are close to one another, say within the interval $[1 - \epsilon, 1 + \epsilon]$, and so the fitness of genotype g can be written as $F_g = 1 + \epsilon\Delta_g$, where ϵ is the selection strength, assumed here to be small, and $\Delta_g \in [-1, 1]$ can be called the differential fitness of the genotype. With this in mind, the equation above can be written

$$x_i^{t+1}(j) = \frac{1}{X^t} x_i^t(j) (1 + \epsilon\Delta_i^t(j)), \quad (1)$$

where $\Delta_i^t(j)$ is the expected differential fitness amongst genotypes that contain allele j at locus i .

We now review the framework of game theory.

A game has several players, and each player i has a set A_i of possible actions. Each player also has a utility, capturing the way whereby her actions and the actions of the other players affect this player's well being. Formally the utility of a player is a function that maps each combination of actions by the players to a real number (intuitively denoting the player's gain, in some monetary unit, if all players choose these particular actions). In general, rather than choosing a single action, a player may instead choose a *mixed* or *randomized* action, that is, a probabilistic distribution over her action set. Here we only need to consider coordination games, in which all players have the same utility function — that is, the interests of the players are perfectly aligned, and their only challenge is to coordinate their choices effectively. Coordination games are among the simplest games; the only challenge in such a game is for the players to “agree” on a mutually beneficial action.

How do the players choose and adjust their choice of randomized (mixed) actions

over repeated play? Assume that at time t , player i has mixed action x_i^t , assigning to each action $j \in A_i$ the probability $x_i^t(j)$. The MWUA algorithm [6] adjusts the mixed strategy for player i in the next round of the game according to the following rule:

$$x_i^{t+1}(j) = \frac{1}{Z^t} x_i^t(j) (1 + \epsilon u_i^t(j)), \quad (2)$$

where Z^t is a normalizing constant designed to ensure that $\sum_j x_i^t(j) = 1$, so x_i^{t+1} is a probability distribution; ϵ is a crucial small positive parameter, and $u_i^t(j)$ denotes the expected utility gained by player i choosing action j in the regime of the mixed actions by the other players effective at time t . This algorithm (a) is known to converge to the min-max actions if the game is two-player zero-sum; (b) is also shown here to converge to equilibrium for the coordination games of interest in the present paper (see Corollary 5 in SI); (c) is a general “learning algorithm” that has been shown to be very successful in both theory and practice; and (d) if, instead of games, it is applied to a large variety of optimization problems, including linear programming, convex programming, and network congestion, it provably converges to the optimum quite fast.

It can be now checked that the two processes expressed in equations (1) and (2), evolution under natural selection in the presence of sex and multiplicative weight updates in a coordination game, are mathematically identical (see Theorem 3 in the SI). That is, the interaction of weak selection and sex is equivalent to the MWUA in a coordination game between loci in which the common utility is the differential fitness of the organism. The parameter ϵ in the algorithm, which, when small signifies that the algorithm is taking a “longer-term view” of the process to be solved (see SI), corresponds to the selection strength in evolution, i.e., the magnitude of the differences between the fitness of various genotypes.

The MWUA is known in computer science as an extremely simple and yet unexpectedly successful algorithm, which has surprised us time and again by its prowess in solving sophisticated computational problems such as congestion minimization in networks and convex programming in optimization. The observation that multiplicative weight updates in a coordination game is equivalent to evolution under sex and

weak selection makes an informative triple connection between three theoretical fields: evolutionary theory, game theory, and the theory of algorithms/machine learning.

So far we have presented the MWUA by “how it works” (informally, it boosts alleles proportionally to how well they do in the current mix). There is an alternative way of understanding the MWUA in terms of “what it is optimizing.” That is, we imagine that the allele frequencies of each locus in each generation is the result of a deliberate optimization by the locus of some quantity, and we wish to determine that quantity.

Returning to the game formulation, define $U_i^t(j) = \sum_{\tau=0}^t u_i^\tau(j)$ to be the cumulative utility obtained by player i by playing strategy j over all t first repetitions of the game, and consider the quantity

$$\sum_j x_i^t(j) U_i^t(j) - \frac{1}{\epsilon} \sum_j x_i^t(j) \ln x_i^t(j). \quad (3)$$

The first term is the current (at time t) expected cumulative utility. The second term of (3) is the *entropy* (expected negative logarithm) of the probability distribution $\{x_i(j), j = 1, \dots, |A_i|\}$, multiplied by a large constant $\frac{1}{\epsilon}$. Suppose now that player i wished to choose the probabilities of actions $x_i^t(j)$'s with the sole goal of maximizing the quantity (3). This is a relatively easy optimization problem, because the quantity (3) to be maximized is strictly concave, and therefore it has a unique maximum, obtained through the KKT conditions [19] (see Section 4 of the SI):

$$U_i^t(j) - \frac{1}{\epsilon}(1 + \ln x_i^t(j)) + \mu^t = 0.$$

(Here μ^t is the Lagrange multiplier associated with the constraint $\sum_j x_i^t(j) = 1$ seeking to keep the $x_i^t(j)$'s a probability distribution, see the SI.) Subtracting this equation from its homologue with t replaced by $t + 1$, and applying the approximation $\exp(\epsilon u_i^t(j)) \approx (1 + \epsilon u_i^t(j))$, we obtain the precise equation (2) (the normalization Z^t is obtained from μ^t and μ^{t+1} , see the SI for the more detailed derivation).

Thus, since equations (1) and (2) are identical, we conclude that, in the weak selection regime, natural selection is tantamount to each locus choosing at each generation

its allele frequencies in the population so as to maximize the sum of the expected cumulative differential fitness over the alleles, plus the distribution's entropy.

This alternative view of selection provides a new insight into an important question in evolutionary biology, namely: How is genetic diversity maintained in the presence of natural selection [81]? That the MWUA process enhances the entropy of the alleles' distribution (while at the same time optimizes expected cumulative utility) hints at such a mechanism. In fact, entropy is enhanced inversely proportional to s (the quantity corresponding in the population genetics domain to the parameter ϵ), the selection strength: The weaker the selection, the more it favors high entropy. Naturally, entropy will eventually vanish when the process quiesces at equilibrium: one allele per locus *will* eventually be fixed, and in fact this equilibrium may be a local, as opposed to global, fitness maximum. However, we believe that it is interesting and significant that the entropy of the allele distribution is favored by selection in the transient; in any event, mutations, environmental changes, and finite population effects are likely to change the process before equilibrium is reached. This new way of understanding the maintenance of variation in evolution (selection as a trade-off between fitness and entropy maximization) is quite different from previous hypotheses for the maintenance of variation (e.g., [40, 75]). Another rather surprising consequence of this characterization is that, under weak selection, all past generations, no matter how distant, have equal influence on the change in the allele mix of the current generation.

Mixability and Diversity

Why is mixability advantageous? (Since sex is ubiquitous, there must be an advantage.) One intuitive explanation may be that, by promoting “good mixer genes” (a phrase by Kimura [32] who had anticipated some of this thinking), it enhances genetic diversity. But what happens at equilibrium (when the allele frequencies finally stabilize, as predicted by our results)?¹ It would be disappointing — and, by the above intuition, detrimental to the argument in favor of mixability — if it so happens that, in the end,

¹In actual Evolution, of course, nothing ever stabilizes, as new mutations introduce new strategies in the game, and life goes on. Note that our model does not include mutations.

natural selection by sex ends up putting all its chips on a single allele per gene (as it often happened in the simulations in [83]). The question arises: *How large is the typical support of an equilibrium for our population genetics process?*

we answer this question in a positive way: there *are* exponentially many equilibria whose support contains a significant fraction of the alleles of each gene, see Corollary 4. We do this through a digression to a completely different problem: Let A be an $n \times n$ matrix. What are the chances that the solution to the system $Ax = 1$ is positive? Assume that the entries of the matrix are iid from a distribution that is continuous and symmetric around zero, say uniform in $[-1, 1]$; in this case, with probability one the solution exists and has no zero term. Intuitively, the answer is 2^{-n} (each row of the inverse must have a positive sum). It is not hard to show that this insight is correct.

But suppose that we want both systems $Ax = 1$ and $A^T y = 1$ to have positive solutions. What are the chances of that happening? One expects this to be about 2^{-2n} , but there is dependence now and the calculation is not straightforward. Intuitively the dependence is favorable, but how does one establish this?

We prove that this probability is *at least* 2^{-2n+1} (and thus the dependence *is* indeed favorable, see Theorem 2). The proof uses a potential function argument reminiscent of the Berlekamp switching game [46].

The connection to Evolution is the following: First, one can show that only square submatrices of the fitness matrix are likely to support an equilibrium of the population dynamics. A square submatrix of the fitness matrix is the support an equilibrium if and only if the corresponding submatrix of *selections* (fitness minus one) has the property that the solutions to both row and column linear systems with unit right-hand sides are positive. By showing that this happens with sufficiently high probability we establish that, in expectation at least, there are equilibria with substantial supports (in fact, quite a few of them), and thus diversity is not always lost at equilibrium. However, note that recent results [94] have shown that the diverse equilibria are for the most part unstable equilibria. This means that diversity in such models is very fragile.

Now let's discuss the proof. We know that the population genetic dynamics converges, but what is the nature of the equilibria? In particular, are they likely to have

extensive support? This would imply that diversity is not totally lost in the process, and considerably strengthen our model as a candidate for the role of sex in Evolution.

To tackle this problem, we need a probabilistic model on the fitness landscape. As we are assuming weak selection, we postulate that the entries of W are drawn iid from a continuous distribution on $[1 - s, 1 + s]$ with no singularities *that is symmetric around 1*. Equivalently, the entries of Δ are iid on $[-1, 1]$. (In fact, our results do not require that the distributions of the entries be identical.)

At equilibrium, all alleles of a gene must have the same mixability (expected fitness with respect to the frequencies of the other alleles). Focusing on the two gene case (here by necessity, because the more general case seems intractable), it must be that $Wx = a\mathbf{1}$ and $W^T y = b\mathbf{1}$ for some real vectors with nonnegative coefficients x, y and some reals a, b (in fact, it is easy to see that a will be equal to b). An equilibrium is characterized by the supports of x and y , or, equivalently, by the submatrix defined by these. Let us call a submatrix A of W an *equilibrium* if $Ax = a\mathbf{1}$ and $A^T y = a\mathbf{1}$ have nonnegative solutions x, y adding to one, for some $a > 1$. We require that $a > 1$ for this reason: If $a < 1$, then A is indeed an equilibrium, but one that leads to extinction.

Consider an $m \times n$ submatrix A of the fitness matrix. When is A an equilibrium? First of all, if A is not square, say $m < n$, then the probability of A being an equilibrium is zero, because then the system $A^T y = a\mathbf{1}$ is overdetermined. So, we shall focus on a square submatrix A . Under weak selection, we can write $A = U + sB$.

Lemma 1. *A is an equilibrium if and only if B^{-1} exists and has positive row and column sums.*

Proof. If A is an equilibrium with $x, y > 0$ the solutions to the linear systems with right-hand sides $a\mathbf{1}$, then it is easy to see that $Bx = \mathbf{1}(a - \sum_j y_j) = \mathbf{1}(a - 1)$ and similarly $B^T y = \mathbf{1}(a - 1)$. Therefore B^{-1} must exist and have positive row and column sums. And from any nonnegative solutions of $Bz = \mathbf{1}, B^T w = \mathbf{1}$ we can get back the solutions of $Ax = a\mathbf{1}, A^T y = a\mathbf{1}, \sum_i x_i = 1, \sum_j y_j = 1$, adding up to one: $x = \frac{z}{\sum_i z_i}$ and similarly $y = \frac{w}{\sum_j w_j}$ with $a = 1 + \frac{s}{\sum_i z_i} > 1$. \square

Thus, to show that there are enough equilibria with large support, we must calculate

(more precisely, lower bound) the probability that a random matrix B has an inverse with positive row and column sums. Our main result is the following:

Theorem 2. *The probability that A is an equilibrium is at least $2^{-(2n-1)}$.*

Proof. By the Lemma, we must bound from below the probability that B^{-1} has positive sums (since B^{-1} exists with probability one).

Let $S \subseteq [n]$. By I_S we denote the $n \times n$ unit matrix with the i th one replaced by -1 whenever $i \in S$. Notice that $I_S E I_T$ is E with all columns in S flipped (multiplied by -1) and all rows in T also flipped (with entries at the intersection of a row in T and a column in S restored to its original value). Notice that $(I_S E I_T)^{-1} = I_T E^{-1} I_S$. That is, to invert E with some rows and columns flipped, you invert E and then flip the same columns and rows, with the roles of columns and rows interchanged.

Now take B and consider all possible flippings $I_S B I_T$. There are 2^{2n-1} distinct such matrices, because it is easy to see that $I_S B I_T = I_{[n]-S} B I_{[n]-T}$ and that all other pairs of flippings are distinct. We shall argue that one of these flippings must have positive row and column sums.

Lemma 3. *For every $B \in [-1, 1]^{n \times n}$ there are $S, T \subseteq [n]$ such that $I_S B I_T$ has non-negative row and column sums.*

Proof. To prove the lemma, start with B and perform the following:

while there is a row or column with negative sum, flip it.

Naturally, after each such flipping other columns or rows, which had positive sums, may become negative. However, if the sum of the row or column that was flipped was $-\sigma$, notice two things: First, $\sigma \geq \sigma_0$, where σ_0 is a constant depending on B ; and second, *the total sum of the entries of E increases by 2σ after the flip*. Therefore, the process must end, and this can only happen if the matrix has positive row and column sums. □

The theorem now follows: Consider the domain $M = [-1, +1]^{n \times n}$, and the subset M_+ whose inverse exists and has positive sums. This subset can be defined using

polynomial inequalities, and is this measurable. By applying the 2^{2n-1} transformations $B \mapsto I_S B I_T$ to M_+ , by the Lemma we exhaust all of M . Therefore, the probability that a matrix in M is in M_+ is at least $2^{-(2n-1)}$. \square

Corollary 4. *The expected number of $k \times k$ equilibria in an $m \times n$ weak selection fitness matrix is at least $2 \left(\frac{mn}{4k^2}\right)^k$.*

Notice that, for k less than half times the geometric mean of m and n , this is exponential in k .

Proof. Use the approximation $\binom{n}{k} \geq \left(\frac{n}{k}\right)^k$. \square

4.2 Supplementary Information

In this section,

- we introduce the detailed modeling assumptions and dynamical model we use for recombination under selection (Section 4.3);
- we introduce the *experts problem* from Computer Science, and the multiplicative weights update algorithm used to solve it (MWUA)
- we point out that the MWUA can be interpreted as an algorithm which at each step optimizes a convex combination of the cumulative sum of gains and of the distribution's entropy;
- we prove the equivalence of the dynamics under weak selection to the MWUA in Theorem 8 with $\epsilon = s$, the selection strength.

Note that some of this material is repeated from the MWU chapter, but it was necessary to re-emphasize in a self-contained way for ease of understanding.

4.3 Population Genetics under selection and recombination

We make several (more or less standard) simplifying assumptions, which are generally trusted not to change substantially the essence of the evolutionary dynamics. The

population of genotypes is infinite. We assume that the genotypes are *haploid* (contain only one copy of each gene), and that the organisms mate at random to produce a new generation; further, we assume there is no overlap between generations (as if all mating happens simultaneously and soon before death). Each offspring's genome is formed by picking, for each gene, an allele from one of the two parent genomes, independently and with probability half each.

Our exposition will be for the case of two genes with m and n alleles respectively, even though our results can be easily seen to extend to any number of genes. Thus genotypes are pairs ij . Each genotype ij has a *fitness value* w_{ij} which is the expected number of offspring the genotype produces (by mating randomly). The matrix $W = [w_{ij}]$, often called the *fitness landscape* of the species, entails the basic genetic parameters of the species (it is a k -dimensional tensor for k genes).

We shall be interested in the statistics of the genotypes in the population. The frequency of the genotype ij will be denoted p_{ij} . The matrix of the p_{ij} 's is the *state* of the dynamical system we shall follow. We denote the value of p_{ij} in generation t by p_{ij}^t .

How do the p_{ij}^t 's change from one generation to the next? Each pair of genotypes mates with a probability determined by the frequencies of those genotypes and recombines with probability $r \in [0, 1/2]$ to produce an offspring, which is then selected. Accordingly, the expected frequency of the genotype ij at the next generation p_{ij}^{t+1} , can be written:

$$p_{ij}^{t+1} = \frac{w_{ij}}{\bar{w}_t} ((1-r)p_{ij}^t + r \sum_l p_{il}^t \sum_k p_{kj}^t)$$

where \bar{w}_t is the sum of the numerators for all ij , so that frequencies add up to 1.

Wright Manifold, Weak Selection, and Nagylaki's Theorem

Besides the p_{ij} frequencies, one has the marginal frequencies, one for each allele: $x_i = \sum_j p_{ij}$ and $y_j = \sum_i p_{ij}$. Within the simplex of the p_{ij} 's, of particular interest to us is the *Wright manifold* on which p_{ij} is a product distribution (the matrix p_{ij} has rank one): $p_{ij} = x_i \cdot y_j$. It turns out that, on the Wright manifold, the population genetic equations take a much simpler form, expressed in terms of the marginal probabilities

x_i and y_j (see Lemma 7).

Life, in general, does not reside on the Wright manifold — that is to say, genotype frequencies do not in general have rank one. This is called *linkage disequilibrium*, and is measured by the distance from the Wright manifold $D_{ij} = p_{ij} - x_i \cdot y_j$. Intuitively, it comes about because differences in the fitness of genotypes distort the allele statistics; just imagine two alleles of two genes whose combination is deleterious. By definition, D_{ij} is zero on the Wright manifold.

Weak selection is an important point of view on evolution, which postulates that the entries of the tensor W are all very close to one another relative to recombination. Differences in fitness are minuscule, and the w_{ij} 's all lie within the interval $[1 - s, 1 + s]$ for some very small $s > 0$ which we call the *selection strength*. It is a mathematical embodiment of the *neutral theory* of Kimura [74], stating roughly that Evolution proceeds mostly through statistical drift due to sampling error that has no impact on fitness.

There is an important connection between the Wright manifold and weak selection, best articulated through Nagylaki's Theorem. Consider the evolution of genome frequencies p_{ij}^t (or for more than two genes) in a situation in which the fitness values are within $[1 - s, 1 + s]$ for some tiny $s > 0$ — that is, weak selection prevails. Consider also the corresponding time series of linkage disequilibria $D_{ij}^t = p_{ij}^t - x_i \cdot y_j$.

Theorem 5. (Nagylaki [100, 103]) (1) for any $t \geq t_0 = 3 \log \frac{1}{s}$ and any i, j , $D_{ij}^t = O(s)$; and furthermore

(2) for $t \geq t_0$ there is a corresponding process $\{\hat{p}_{ij}^t\}$ on the Wright manifold such that (a) $|\hat{p}_{ij}^t - p_{ij}^t| = O(s)$; and (b) both processes converge and there is one-to-one correspondence between the equilibria of p_{ij}^t and the equilibria of \hat{p}_{ij}^t .

Nagylaki's Theorem states essentially that, in order to understand a genotype frequency process in the weak selection regime, one can instead follow a closely related process on the Wright manifold. As we shall see next, it turns out that this brings about some unexpected connections.

4.4 The Experts Problem

We now discuss a seemingly completely unrelated problem from Computer Science, and an important algorithm used to solve it.

Imagine that every day you receive advice from n financial experts, and then you must select one of them and follow his advice for that day. Following the advice of expert i in day t results in a net gain (or loss) of g_i^t , a number between -1 and 1 . The g_i^t 's are arbitrary numbers in this range, and are not known in advance. This process is repeated for a large number T of days. In the end of the T days, the optimum expert is the one with the largest cumulative gain $G_i = \sum_{t=1}^T g_i^t$; let i^* be this expert, and G^* be this maximum cumulative gain. We wish to come up with an algorithm — possibly randomized — for selecting an expert on each day so that in the end of the T days our total gain is in expectation very close to G^* . In other words, we want to achieve, in the end of the T days, a performance very close to the performance of the expert who is best *in retrospect*, even though the g_i^t 's are unknown and arbitrary — for example, they could be chosen by an adversary striving to deteriorate the performance of the algorithm.

This ambitious goal (which, some would argue, seems intuitively impossible to achieve) can be attained by a very simple method called *multiplicative weight update algorithm* or MWUA. This method has been first discovered in by the economist J. Hannan in connection with repeated games [56], then rediscovered by Thomas Cover in relation to portfolio analysis [29], later it was used in Artificial Intelligence under the name “Boosting [51], and earlier in a version called “Winnow” [82], until it was recognized as the common idea underlying several simple and curiously effective optimization algorithms developed by computer scientists to solve linear and convex programming problems and network congestion problems, among many others, and codified as MWUA [6].

The MWUA assigns each day t *weights*, or probabilities, $p_i^t > 0$ to the experts, and each day selects at an expert at random among the n with these probabilities. Initially all weights are, say, equal $p_i^1 = \frac{1}{n}$ for all i . Then each day, the weight of the i th expert

is modified as follows:

$$p_i^{t+1} = \frac{1}{\bar{Z}^{t+1}} p_i^t (1 + \epsilon g_i^t) \quad (MWUA)$$

where the normalization $\bar{Z}^{t+1} = \sum_{i=1}^n p_i^t (1 + \epsilon g_i^t)$ keeps the weights probabilities, and $\epsilon > 0$ is a small constant chosen to balance long-term risk with short-term gains (in the experts problem, a good choice of ϵ turns out to be $\sqrt{\frac{\ln n}{T}}$). That is, the probability of selecting expert i is “boosted” by a small multiple of the expert’s gain that day (decreased if that gain is negative).

The following result now captures the surprisingly favorable performance of this simple algorithm:

Theorem 6. *The total gain achieved by the MWUA is in expectation at least $(1 - \epsilon) \cdot G^* - \frac{\ln n}{\epsilon}$.*

To see how favorable is the performance of MWUA as stated by this result, notice that it comes ϵ close to the optimum, minus a quantity that does not depend on T . To put it differently, if we choose $\epsilon = \sqrt{\frac{\ln n}{T}}$, on an average day this algorithm does only $O(\frac{1}{\sqrt{T}})$ worse than the ex post best expert.

4.5 An Optimization Interpretation of MWUA

In this section we point out that the MWUA can be thought of as a multi-step optimization algorithm which, at each step, strives to optimize a convex combination of (1) the expected cumulative gain; and (2) the entropy of the experts’ distribution. This interpretation is implicit in the literature [136, 109].

Let $G_i^t = \sum_{\tau=1}^t g_i^\tau$ be the cumulative gain of expert i in the first t days; thus, $G^* = \max\{G_1^T, \dots, G_n^T\}$. Suppose now that at step t we wish to choose the probabilities p_i^t so as to maximize the sum of two quantities: The expected cumulative gain so far, and (some positive constant $\frac{1}{\epsilon}$ times) the *entropy* of the distribution p_i^t ’s, $-\sum_{i=1}^n p_i^t \ln p_i^t$. That is,

$$\max_{\{p_i^t\}} \sum_{i=1}^n p_i^t G_i^t - \frac{1}{\epsilon} \sum_{i=1}^n p_i^t \ln p_i^t,$$

subject to $\sum_i p_i^t = 1$. Now this is a strictly convex optimization problem, as one term

is linear and the other strictly concave, and thus it has a unique optimum, which can be found through the KKT-conditions [19]:

$$G_i^t - (1 + \ln p_i^t) + \mu^t = 0, i = 1, \dots, n,$$

where μ^t is the Lagrange multiplier corresponding to the equality constraint. Similarly, we can write the same equation for the next generation, with t replaced by $t + 1$:

$$G_i^{t+1} - (1 + \ln p_i^{t+1}) + \mu^{t+1} = 0, i = 1, \dots, n,$$

Subtracting these two equations and solving for p_i^{t+1} , and recalling that $G_i^{t+1} - G_i^t = g_i^{t+1}$, we obtain precisely equation (MWUA), where the normalization $\bar{Z}^{t+1} = \exp(\mu^{t+1} - \mu^t)$.

4.6 Coordination Games between Genes

We now introduce the basic formalism of Game Theory. In a *game* each of finitely many *players* has a set of *strategies*, and a *payoff function* mapping the cartesian product of the strategy sets to the reals. A game in which all payoff functions are identical is called a *coordination game*. In a coordination game the interests of all players are perfectly aligned, and, intuitively, they all strive to hit the same high value of the common payoff function. In terms of equilibrium calculations, they are trivial.

Fix a game, and a mixed strategy profile, that is, for each player p a distribution x^p over her strategies. For each player p and each strategy $a \in S_p$ one can calculate the expected payoff of this strategy, call it $q(a)$. How does the player's strategy change in time? One possibility is inspired by the MWUA of the previous section. The *multiplicative weight update dynamics* of the game transform the mixed strategy profile $\{x^p\}$ as follows: For each player p and each strategy $a \in S_p$, the probability $x^p(a)$ of player p playing a becomes

$$\frac{x^p(a) \cdot (1 + \epsilon \cdot q(a))}{1 + \epsilon \cdot \sum_{b \in S_p} x^p(b)q(b)} = \frac{x^p(a) \cdot (1 + \epsilon \cdot q(a))}{1 + \epsilon \cdot \bar{q}},$$

where by \bar{q} we denote the expected payoff to p (in a coordination game, to all players). That is, the probability of playing a is boosted by an amount proportional to its expected payoff, and then renormalized. It is known that two players following the multiplicative update dynamics attain the Nash equilibrium in zero-sum games (this has been rediscovered many times over the past fifty years, see for example [6]), but not in general games. It follows directly from the results below that it also converges to the Nash equilibrium in coordination games. Beyond games and portfolio management, the multiplicative updates dynamics lies at the foundations of a very simple, intuitive, robust and powerful algorithmic idea of very broad applicability [6].

Going now back to population genetics dynamics, let w_{ij} be a fitness landscape (matrix for two genes, tensor for more) in the weak selection regime, that is, each entry is in the interval $[1 - s, 1 + s]$. Define the *differential fitness landscape* to be the tensor with entries $\Delta_{ij} = \frac{w_{ij} - 1}{s}$.

We next point out a useful way to express the important analytical simplification afforded by the Wright manifold:

Lemma 7.) *On the Wright manifold, the population genetics dynamics becomes*

$$p_{ij}^{t+1} = \frac{1}{\bar{w}_t} x_i^t \cdot y_j^t \cdot w_{ij},$$

and similarly for more genes.

Proof. As is shown in ref. [83], we can re-write the population genetics dynamics as:

$$p_{ij}^{t+1} = \frac{1}{\bar{w}_t} w_{ij} (p_{ij}^t - r D_{ij}^t)$$

where $D_{ij}^t = p_{ij}^t - x_i^t y_i^t$ is the linkage disequilibrium and \bar{w}_t is a normalization constant.

Now because $D_{ij} = 0$, we have:

$$p_{ij}^{t+1} = \frac{1}{\bar{w}_t} w_{ij} (p_{ij}^t)$$

Finally, because $D_{ij}^t = 0$, $p_{ij}^t = x_i^t y_j^t$. The result follows. \square

We are now ready for the main result of this section:

Theorem 8. *Under weak selection with selection strength s , the population genetic dynamics is precisely the multiplicative update dynamics of a coordination game whose payoff matrix is the differential fitness landscape and $\epsilon = s$.*

Proof. We only show the derivation for two genes, the general case being a straightforward generalization.

$$x_i^{t+1} = \sum_j p_{ij}^{t+1} = \frac{1}{\bar{w}_t} \sum_j x_i^t y_j^t w_{ij} = \frac{x_i^t}{\bar{w}_t} (1 + s \sum_j y_j^t \Delta_{ij}) = \frac{x_i^t \cdot (1 + s \sum_j y_j^t \Delta_{ij})}{1 + s \cdot \Delta}.$$

Here the first equation is the definition of marginal frequencies, the second is the Lemma, the third uses the definition of Δ_{ij} , and the last one follows from the expectation of w_{ij} being 1 plus s times the expectation of Δ_{ij} . The last expression is precisely the multiplicative update dynamics, completing the proof. \square

Finally, we can connect our result to the optimization interpretation of MWUA:

Corollary 9. *Under weak selection with selection strength s , the population genetics dynamics is tantamount to each gene optimizing at generation t a quantity equal to the cumulative expected fitness over all generations up to t , plus $\frac{1}{s}$ times the negative entropy of the allele distribution of the gene at time t .*

One interpretation is this: If the optimization of cumulative expected fitness is sought, then it makes sense at each step, and in view of the uncertainties of future steps, to balance off cumulative expected fitness so far against the distribution's entropy (a well-known measure of dispersion).

Mixability and MWUA

The mathematical theory of population genetics has long attempted to gain a deep understanding of population genetic dynamics in the presence of natural selection and thus of how evolution works. However, the common mode of reproduction in nature is

the sexual one and, from the start, understanding the population genetic dynamics in the presence of both sex and natural selection has been a complicated task [45, 28].

Mixability theory [83, 86, 85] examines the population genetic dynamics in the presence of sex and natural selection using the classical population genetic equations. In that, it follows a century of population genetics research. However, unlike other theories, it focuses on how well alleles perform across a variety of genetic contexts. It holds that in the presence of sex, this ability, called “mixability,” is increased [83, 86].

Aside of a brief comment on mixability by Crow and Kimura [32] and a related simulation study [96], this simple idea has not been part of the mathematical investigation of the basis of evolution. This has been at least partly due to a technical reason [83]. Selection for mixability occurs when there is genetic diversity across loci, so that alleles in one gene can be selected based on how well they perform with a diversity of alleles at other genes. This means that it occurs during the evolutionary transient [83]. While evolution happens in the transient, mathematical models often examine equilibria, leaving the less wieldy transient relatively little studied [83].

The first breakthrough in the theoretical study of mixability was due to computer iterations and the definition of a new mathematical measure that made it possible to capture selection for mixability and observe it clearly during the evolutionary transient [83]. Later, Livnat et al. proposed the first theorem of mixability [85]; however, it was very limited in applicability because it relied on unrealistic conditions for the purposes of mathematical tractability [85]. Our purpose here is to bring to the attention of the theoretical biological community as well as clarify a far more powerful result, capturing analytically and with generality the fact that, during the transient, mixability is favored.

This result is based on a newly founded bridge between theoretical computer science and evolutionary biology. Recently, we have shown that, in the realm of weak selection, using a haploid infinite population model, the multilocus population genetic dynamics in the presence of sex and natural selection can be described as repeated play of a coordination game, where the genes are the players, the alleles are the strategies, allele frequencies are the probabilities in the mixed strategies, the utility is the organismal fitness, and furthermore, the game is played according to a powerful algorithm, known

in computer science as the Multiplicative Weights Update Algorithm (MWUA) [25]. This new bridge allows us to use results and techniques known in theoretical computer science—especially here the no-regret theorem of Arora et al. [6]—in evolutionary theory. In particular, we show that in the presence of natural selection and sex, the time-averaged population mean fitness is greater or equal to the largest time-averaged allelic mixability, averaged over loci, minus small terms. This general result means that mixability is an important guiding principle with which to understand population genetic dynamics, one which may in turn affect our understanding of the population mean fitness and more.

Model and theorem

In Chastain et al. [25], a haploid model with two genes was considered, with n_1 and n_2 alleles in the first and second loci respectively, where genotype ij consists of alleles $i \in \{1, \dots, n_1\}$ and $j \in \{1, \dots, n_2\}$; x_i^t and y_j^t are the frequencies of alleles i and j at generation t ; w_{ij} and q_{ij}^t are the fitness and the frequency of genotype ij at generation t , respectively; and \bar{w}_t is the population mean fitness at generation t , $\sum_{i,j} w_{ij} q_{ij}^t$. Furthermore, the fitness values were taken to be within the interval $[1 - s, 1 + s]$ where s is small, and Δ_{ij} was defined as follows: $\Delta_{ij} = (w_{ij} - 1)/s$.

In this paper, we will operationalize mixability—the ability of an allele to perform well across contexts—as the expected fitness of the allele, namely $\sum_{j=1}^{n_2} y_j^t w_{ij}$ for allele i at the first locus and $\sum_{i=1}^{n_1} x_i^t w_{ij}$ for allele j at the second locus. Using Chastain et al.’s model, we will prove the following:

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} x_i^t y_j^t w_{ij} &\geq \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\sum_{j=1}^{n_2} y_j^t w_{ij} + \sum_{i=1}^{n_1} x_i^t w_{ij'} \right) \\ &\quad - s^2 \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\left| \sum_{j=1}^{n_2} y_j^t \Delta_{ij} \right| + \left| \sum_{i=1}^{n_1} x_i^t \Delta_{ij'} \right| \right) - \frac{1}{2T} \ln(n_1 + n_2), \forall i', j'. \end{aligned}$$

In words: the time-averaged population mean fitness up to generation T (on the left) is greater or equal to the time-averaged mixability (in the sense of expected fitness) of

any allele, one for each locus, averaged over loci, minus small terms. In particular, this is true for the best (most mixable) alleles. Since the population mean fitness cannot be larger than the mixability of the maximal-mixability allele at either locus, this implies that the performance of the population as a whole is close to the performance of the most (cumulatively) mixable alleles.

Proof. We first provide background on the Multiplicative Weights Update Algorithm (MWUA) [6], and then adjust the MWUA no-regret theorem [6] to the population genetic context.

Arora et al. [6] present the MWUA as follows. Assume that n experts provide advice at each of T time periods. Following the advice of expert $i \in n$ at time period $t \in T$ will yield the gain $m_i^{(t)} \in [-1, 1]$. (For example, one can think of these experts as investment options, each of which can gain or lose up to a certain percentage of the money investment in them.) The goal is to find an advice-following or investment strategy such that, after many time periods, the strategy will perform almost as well as a strategy that would have consistently followed the advice of the expert found to be best overall in retrospect, even though the gains of the experts are not known in advance (no assumption is made about the experts performance—the gains may be correlated, they may depend on past choices, and the experts are even allowed to act adversarially).

At first sight, such a strategy seems impossible. But the MWUA in fact achieves it. The MWUA assigns a weight w_i to each expert i , being 1 for each expert at the first time period. At each time period, the expert whose advice is to be followed is chosen with a probability proportional to its weight. The vector of advice-following probabilities is then $p^{(t)} = \{w_1^{(t)}/\Phi^{(t)}, \dots, w_n^{(t)}/\Phi^{(t)}\}$, where $\Phi^{(t)} = \sum_i w_i^{(t)}$. The gains of the different experts are then observed, and the weights are updated according to the following rule: $w_i^{(t+1)} = w_i^{(t)}(1 + \eta m_i^{(t)})$, using a fixed $\eta \leq \frac{1}{2}$.

Theorem 2.5 in Arora et al. states that, under the MWUA, for any agent i :

$$\sum_{t=1}^T m^{(t)} \cdot p^{(t)} \geq \sum_{t=1}^T m_i^{(t)} - \eta \sum_{t=1}^T |m_i^{(t)}| - \frac{\ln n}{\eta}.$$

For large T , the last term becomes negligible, and for small η , the second term is small. Now, since this inequality holds for any expert i , it holds specifically also for the best expert in retrospect, and thus the performance of the MWUA is close to that of the best expert in retrospect.

As shown in Chastain et al. (see proof of Theorem 4),

$$x_i^{t+1} = \frac{1}{\bar{w}_t} \sum_j x_i^t y_j^t w_{ij} = \frac{x_i^t}{\bar{w}_t} \sum_j y_j^t w_{ij} = \frac{x_i^t}{\bar{w}_t} \sum_j y_j^t (1 + \Delta_{ij}s) = \frac{x_i^t}{\bar{w}_t} \left(1 + s \sum_j y_j^t \Delta_{ij} \right).$$

Now, replace η with s , p_i with x_i (same for y_i), and m_i with $d_1^t(i)$, where

$$d_1^t(i) = \sum_{j=1}^{n_2} y_j^t \Delta_{ij}.$$

Using Theorem 2.5 in Arora et al., the following holds:

$$\sum_{t=1}^T \sum_{i=1}^{n_1} d_1^t(i) x_i^t \geq \sum_{t=1}^T d_1^t(i') - s \sum_{t=1}^T |d_1^t(i')| - \frac{\ln(n_1)}{s}, \forall i'.$$

Using the same inequality for the second locus (with $d_2^t(j) = \sum_{i=1}^{n_1} x_i^t \Delta_{ij}$), we get:

$$\begin{aligned} \sum_{t=1}^T \left(\sum_{i=1}^{n_1} d_1^t(i) x_i^t + \sum_{j=1}^{n_2} d_2^t(j) y_j^t \right) &\geq \sum_{t=1}^T (d_1^t(i') + d_2^t(j')) \\ &\quad - s \sum_{t=1}^T (|d_1^t(i')| + |d_2^t(j')|) - \frac{\ln(n_1 + n_2)}{s}, \forall i', j'. \end{aligned}$$

Bringing back the detail, we get

$$\begin{aligned} \sum_{t=1}^T \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} y_j^t \Delta_{ij} x_i^t + \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} x_i^t \Delta_{ij} y_j^t \right) &\geq \sum_{t=1}^T \left(\sum_{j=1}^{n_2} y_j^t \Delta_{i'j} + \sum_{i=1}^{n_1} x_i^t \Delta_{ij'} \right) \\ &\quad - s \sum_{t=1}^T \left(\left| \sum_{j=1}^{n_2} y_j^t \Delta_{i'j} \right| + \left| \sum_{i=1}^{n_1} x_i^t \Delta_{ij'} \right| \right) - \frac{\ln(n_1 + n_2)}{s}, \forall i', j'. \end{aligned}$$

Multiply by $\frac{s}{2T}$ and add 1, and we get

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} x_i^t y_j^t (1 + s\Delta_{ij}) &\geq \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\sum_{j=1}^{n_2} y_j^t (1 + s\Delta_{i'j}) + \sum_{i=1}^{n_1} x_i^t (1 + s\Delta_{ij'}) \right) \\ &\quad - s^2 \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\left| \sum_{j=1}^{n_2} y_j^t \Delta_{i'j} \right| + \left| \sum_{i=1}^{n_1} x_i^t \Delta_{ij'} \right| \right) - \frac{1}{2T} \ln(n_1 + n_2), \forall i', j', \end{aligned}$$

which is the same inequality as the one we started with. \square

Note that the fixed fitness components represented by the Δ_{ij} terms can be replaced with fitness components that vary in time, Δ_{ij}^t , without a change in the proof—the results hold for a constant as well as for a variable environment, and are thus very general, with the exception of the limitation to haploids.

Note that the time-averaged mixability of an allele as defined here constitutes one way of representing the ability of this allele to perform well across a wide variety of genetic combinations. However, note that this mixability measure is not identical to the one used in Livnat et al. [83], where the context of the numerical study necessitated a different measure. Because these two measures both give us a window into the notion of the ability of alleles to perform well across different genetic contexts (and future research may find other such measures), we consider both to be measures of mixability.

There are limits to Nagylaki's theorem, as has been deftly pointed out by [95]. They re-discover regimes in which Nagylaki's theorem does not apply, which he himself points out in the full version of the Theorem [100]. They give examples in which the product distribution is far from the full joint distribution for initial conditions or run-times outside the interval $[t_1, K/s]$ (where $t_1 \sim \log s / \log(1 - r)$ and for some constant $K > 0$) in which Nagylaki's theorem was shown to hold [100]. They also show that when one does selection before recombination, Nagylaki's theorem does not apply. Indeed, Nagylaki's theorem was shown for dynamics in which recombination happens before selection, and our use of it is assuming this kind of dynamics.

In more detail, Equation 53a in Nagylaki[100] gives the conditions necessary for the product distribution to be approximately equal to the full joint distribution (with

an error of $O(s)$). Nagylaki's theorem only holds for times less than K/s and greater than initial time $t_1 \sim \log s / \log(1 - r)$. Nagylaki says: "If $\pi(t)$ does not necessarily converge to some equilibrium point or if $\pi(t_1)$ is on the stable manifold of an unstable equilibrium, then small perturbations may cause large deviations in its ultimate state. In this case, the restriction $t < K/s$ in (53) may be necessary" where $\pi(t)$ is the product distribution. In one example given by Meir & Parkes [95], the time at which divergence happens is well before the initial time t_1 required for Nagylaki's theorem to hold (and by the time it does hold, the product distribution update and the full joint distribution from the resulting initial condition are trivially close). In another example given by Meir & Parkes [95], divergence is shown for dynamics in which selection happens before recombination, for which Nagylaki's theorem does not apply.

4.6.1 Generalizations: Strong selection and Diploid organisms

In the general case of strong selection, we can derive a similar result as for weak selection, but with an extra factor taking into account linkage disequilibrium, which we define as follows:

$$D_{ij}^{(t)} = \frac{p_{ij}^t}{x_i^t y_j^t}$$

The results of Meir and Parkes for the general case of strong selection uses a different generalization of mixability to the strong selection case. We will present both side-by-side and compare. Notably, they don't use the same quantity, linkage disequilibrium, in their definition. But first some preliminary definitions.

We define the conditional probability of allele j arising in the population given allele i arises as:

$$p(j | i)^t = p_{ij}^t / x_i^t$$

$$p(i | j)^t = p_{ij}^t / y_j^t$$

Notice that the linkage disequilibrium can be related to both p_{ij} and $p(j | i)$:

$$D_{ij}^t = p(j | i)^t / y_j^t$$

$$D_{ij}^t = p(i | j)^t / x_i^t$$

For the generalized mixabilities in this setting, we use for allele i , $\sum_j D_{ij}^t y_j^t w_{ij}$ and, for allele j , $\sum_i D_{ij}^t x_i^t w_{ij}$. Recall that the original form for the mixability of i was $\sum_j y_j^t w_{ij}$, and, for allele j , $\sum_i x_i^t w_{ij}$ so our generalization of the mixability is the mixability for the weak-selection case scaled by the linkage disequilibrium. Such a generalization is then a kind of disequilibrium-adjusted mixability. Meir and Parkes use a generalized mixability that is of the form $\sum_j p(j | i)^t w_{ij}$ for allele i and $\sum_i p(i | j)^t w_{ij}$ for allele j . Note that though the two definitions can be shown by probability theory and the identities above to be mathematically equivalent, our definition is different conceptually and turns out to be easier to use mathematically to derive novel results. In particular, using our measure, we are able to prove a novel no-regret result for the model of recombination used here (recombination before selection), and by their mathematical identity, also for Meir and Parkes' mixability measure. Notably, Meir and Parkes showed a no-regret result using their mixability measure for the selection before recombination model, and not for the recombination before selection model. Indeed, Meir and Parkes state that they cannot find this inequality for their mixability measure, and prove a different inequality instead.

Using these identities, we can write the marginal dynamics as follows:

$$x_i^{t+1} = \frac{1}{\bar{w}_t} \sum_j p_{ij}^t w_{ij} = \frac{1}{\bar{w}_t} \sum_j x_i^t p(j | i)^t w_{ij} = \frac{1}{\bar{w}_t} \sum_j x_i^t D_{ij}^t y_j^t w_{ij} = \frac{1}{\bar{w}_t} x_i^t \sum_j D_{ij}^t y_j^t w_{ij}$$

Now for the second locus, allele j :

$$y_j^{t+1} = \frac{1}{\bar{w}_t} \sum_i p_{ij}^t w_{ij} = \frac{1}{\bar{w}_t} \sum_i y_j^t p(i | j)^t w_{ij} = \frac{1}{\bar{w}_t} \sum_i y_j^t D_{ij}^t x_i^t w_{ij} = \frac{1}{\bar{w}_t} y_j^t \sum_i D_{ij}^t x_i^t w_{ij}$$

And using a similar proof as before (with Theorem 2.5), using instead of the mixability, $\sum_j D_{ij}^t y_j^t w_{ij}$ and $\sum_i D_{ij}^t x_i^t w_{ij}$ respectively, we get:

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} x_i^t y_j^t D_{ij}^t (1 + s\Delta_{ij}) \geq \\ & \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\sum_{j=1}^{n_2} D_{ij}^t y_j^t (1 + s\Delta_{i'j}) + \sum_{i=1}^{n_1} D_{ij}^t x_i^t (1 + s\Delta_{ij'}) \right) \\ & - s^2 \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\left| \sum_{j=1}^{n_2} D_{ij}^t y_j^t \Delta_{i'j} \right| + \left| \sum_{i=1}^{n_1} D_{ij}^t x_i^t \Delta_{ij'} \right| \right) - \frac{1}{2T} \ln(n_1 + n_2), \forall i', j', \end{aligned}$$

thus showing that the cumulative product of the mixability and linkage disequilibrium of the population under the dynamics is close to the pair of alleles which maximizes the cumulative product of mixability and linkage disequilibrium.

In the diploid case, we can prove an analogous result. Let us look at the definitions of mixability in the haploid case again. For allele i at the first locus, it is $\sum_{j=1}^{n_2} y_j^t w_{ij}$. For allele i at the first locus, it is $\sum_{i=1}^{n_1} x_i^t w_{ij}$. How would this change in the diploid setting? We propose the following definitions. For the first locus, we define diploid-mixability as:

$$\sum_{j=1}^{n_2} \sum_{kl} y_j^t x_k^t y_l^t w_{ijkl}$$

For the second locus, diploid-mixability is:

$$\sum_{i=1}^{n_1} \sum_{kl} x_i^t x_k^t y_l^t w_{ijkl}$$

We will show that

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{kl} x_i^t y_j^t x_k^t y_l^t (1 + s\Delta_{ijkl}) \geq \\ \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\sum_{jkl} y_j^t x_k^t y_l^t (1 + s\Delta_{i'jkl}) + \sum_{ikl} x_i^t x_k^t y_l^t (1 + s\Delta_{ij'kl}) \right) \\ - s^2 \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\left| \sum_{jkl} y_j^t x_k^t y_l^t \Delta_{i'jkl} \right| + \left| \sum_{ikl} x_i^t x_k^t y_l^t \Delta_{ij'kl} \right| \right) - \frac{1}{2T} \ln(n_1 + n_2), \forall i', j', \end{aligned}$$

We prove this using a similar approach to what was used for the haploid case.

Proof. From the two-locus case of the diploid dynamics [102] (pg. 176 , Eq 8.10 and 8.16):

$$\bar{w}P_{ij}^{t+1} = P_{ij}^t w_{ij} - cD_{ij}^t$$

Where D_{ij}^t are the linkage disequilibria, and $w_{ij} = \sum_{kl} w_{ij,kl} P_{kl}^t$.

Now $D_{ij}^t = 0$ for all i, j on the Wright manifold, so this becomes

$$\bar{w}P_{ij}^{t+1} = P_{ij}^t w_{ij}$$

Now $P_{ij}^t = x_i^t y_j^t$, so

$$\bar{w}P_{ij}^{t+1} = x_i^t y_j^t w_{ij}$$

And thus

$$\bar{w}x_i^{t+1} = \sum_j P_{ij}^{t+1} = \sum_j x_i^t y_j^t w_{ij} = x_i^t \sum_j y_j^t w_{ij}$$

And therefore:

$$\bar{w}x_i^{t+1} = x_i^t \sum_j y_j^t \sum_{kl} w_{ij,kl} P_{kl}^t = x_i^t \sum_j y_j^t \sum_{kl} kl w_{ij,kl} P_{kl}^t$$

$$\bar{w}x_i^{t+1} = x_i^t \sum_j y_j^t (1 + s \sum_{kl} \Delta_{ijkl} P_{kl}^t)$$

Again using Nagylaki's Theorem, we obtain

$$\bar{w}x_i^{t+1} = x_i^t \sum_j y_j^t (1 + s \sum_{kl} \Delta_{ijkl} x_k^t y_l^t)$$

Simplifying,

$$\bar{w}x_i^{t+1} = x_i^t (1 + s \sum_j y_j^t \sum_{kl} \Delta_{ijkl} x_k^t y_l^t)$$

$$\bar{w}x_i^{t+1} = x_i^t (1 + s \sum_{jkl} \Delta_{ijkl} y_j^t x_k^t y_l^t)$$

And also

$$y_j^{t+1} = \sum_i \bar{w} P_{ij}^{t+1} = \sum_i x_i^t y_j^t w_{ij} = y_j^t \sum_i x_i^t w_{ij}$$

And by a similar argument as for x_i^t ,

$$y_j^{t+1} = y_j^t (1 + s \sum_{ikl} \Delta_{ij,kl} x_i^t x_k^t y_l^t)$$

Summarizing the above results, we have for both loci the following:

$$x_i^{t+1} = x_i^t \left(1 + s \sum_{jkl} y_j^t x_k^t y_l^t \Delta_{ijkl} \right)$$

and likewise for the second locus,

$$y_j^{t+1} = y_j^t \left(1 + s \sum_{ikl} x_i^t x_k^t y_l^t \Delta_{ij,kl} \right)$$

Now, as in the haploid case, to use the inequality given by Theorem 2.5 in Arora et al., replace η with s , p_i with x_i (same for y_i), and m_i with $d_1^t(i)$, where

$$d_1^t(i) = \sum_{jkl} y_j^t x_k^t y_l^t \Delta_{ijkl}$$

Using Theorem 2.5 in Arora et al., the following holds:

$$\sum_{t=1}^T \sum_{i=1}^{n_1} d_1^t(i) x_i^t \geq \sum_{t=1}^T d_1^t(i') - s \sum_{t=1}^T |d_1^t(i')| - \frac{\ln(n_1)}{s}, \forall i'.$$

Using the same inequality for the second locus (with $d_2^t(j) = \sum_{ikl} x_i x_k y_l \Delta_{ijkl}$), we get:

$$\begin{aligned} \sum_{t=1}^T \left(\sum_{i=1}^{n_1} d_1^t(i) x_i^t + \sum_{j=1}^{n_2} d_2^t(j) y_j^t \right) \geq \\ \sum_{t=1}^T (d_1^t(i') + d_2^t(j')) - s \sum_{t=1}^T (|d_1^t(i')| + |d_2^t(j')|) - \frac{\ln(n_1 + n_2)}{s}, \forall i', j'. \end{aligned}$$

Bringing back the detail, we get

$$\begin{aligned} \sum_{t=1}^T \left(\sum_{i=1}^{n_1} \sum_{jkl} y_j^t x_k^t y_l^t \Delta_{ijkl} x_i^t + \sum_{j=1}^{n_2} \sum_{ikl} x_i^t x_k^t y_l^t \Delta_{ijkl} y_j^t \right) \geq \\ \sum_{t=1}^T \left(\sum_{jkl} y_j^t x_k^t y_l^t \Delta_{i'jkl} + \sum_{ikl} x_i x_k y_l \Delta_{ij'kl} \right) \\ - s \sum_{t=1}^T \left(\left| \sum_{jkl} y_j^t x_k^t y_l^t \Delta_{i'jkl} \right| + \left| \sum_{ikl} x_i^t x_k^t y_l^t \Delta_{ij'kl} \right| \right) - \frac{\ln(n_1 + n_2)}{s}, \forall i', j'. \end{aligned}$$

Multiply by $\frac{s}{2T}$ and add 1, and we get

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{kl} x_i^t y_j^t x_k^t y_l^t (1 + s\Delta_{ijkl}) \geq \\ \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\sum_{jkl} y_j^t x_k^t y_l^t (1 + s\Delta_{i'jkl}) + \sum_{ikl} x_i^t x_k^t y_l^t (1 + s\Delta_{ij'kl}) \right) \\ - s^2 \frac{1}{T} \sum_{t=1}^T \frac{1}{2} \left(\left| \sum_{jkl} y_j^t x_k^t y_l^t \Delta_{i'jkl} \right| + \left| \sum_{ikl} x_i^t x_k^t y_l^t \Delta_{ij'kl} \right| \right) - \frac{1}{2T} \ln(n_1 + n_2), \forall i', j', \end{aligned}$$

which is the same inequality as the one we started with. \square

In words: the time-averaged population mean cumulative diploid-mixability up to generation T (on the left) is greater or equal to the mean cumulative diploid mixability of any allele, one for each locus, averaged over loci, minus small terms. In particular, this is true for the most (cumulatively) diploid-mixable alleles. Since the population mean fitness cannot be larger than the mixability of the maximal-mixability allele at either locus, this implies that the performance of the population as a whole is close to the performance of the most (cumulatively) diploid-mixable alleles.

Discussion

Although in the asexual case, evolution can be represented as a MWUA process with a single “gene” (representing the entire genome), and in this case is equivalent to replicator dynamics [123], here we have focused on evolution in the presence of recombination, and thus on mixability [83, 86, 85].

Note that, near equilibrium, the result does not tell us much more than we would have expected: as diversity decreases and nearly disappears in this haploid model as the population approaches equilibrium, one allele in each gene comes to be nearly fixed. Then, the population mean fitness is expected to be close to the right-hand side of the last inequality above, for the following reason. If there was an allele of near zero frequency that actually had higher mixability than the dominating allele in its locus during this period of time, we would have expected it to increase in frequency.

However, during the transient, the theorem is informative, and captures analytically and with generality the fact that sex favors mixability. Notice that the last term in the last inequality, $-\frac{1}{2T} \ln(n_1 + n_2)$, rapidly decreases with time. Thus, under constant environment (constant fitness values), the inequality becomes informative after $O(\log n)$ generations and before equilibrium.

The transient may be relevant not only because the number of generations required to reach equilibrium has not passed, but also for any other reason that prevents the dynamics from reaching equilibrium. Environmental change is one such reason. Importantly, one useful aspect of the generality of the present results is that, no matter what temporal change occurs in the fitness values, the last inequality holds. Therefore, under any kind of environmental fluctuation that maintains genetic diversity, our result immediately demonstrates the persistence of mixability.

While Livnat et al.'s [83] goal was to compare the population genetic dynamics with and without sex with respect to mixability, this comparison cannot be done here. With a slight alteration, the present result can be extended beyond weak selection, which means that it holds also when recombination is low and even zero. However, the fact that it applies under any mode of reproduction does not mean that the concept of mixability does. For asexual dynamics, we know *a priori* that the selection only favors the best combination of alleles in a model such as examined here, and any measurement of mixability does not carry with it the same meaning as it does in the sexual case. Only in the sexual case does it imply an increase in the ability of alleles to perform well across different contexts. In other words, the results reported here show that, even in the presence of sexual recombination and complex genetic interactions within and across loci, alleles are favored during the evolutionary transient that perform well with each other.

Finally, our use of the weak selection assumption builds on Nagylaki's theorem, which states that weak selection in the presence of sex proceeds near the Wright manifold [100, 101]. After having extended some but not all of our previous results on population genetics as MWUA [25] by removing the weak-selection assumption, Meir and Parkes argued for the importance of removing this assumption in general [95].

However, since their arguments on this point require specific parameters outside of the range where Nagylaki's theorem holds [100, 101], or specific parameters together with an alternative calculation of the population genetic dynamics not used in [25], in our principal result we have maintained the weak selection assumption, which provides for a simple and clear demonstration of mixability principles. The strong selection generalization of the theorem uses a linkage dis-equilibrium weighed mixability to give a generalized result, in contrast with the conditional probability used in [95] (which has a less clear interpretation in terms of genetics). Nonetheless we acknowledge that their approach to the strong selection case was an inspiration for ours, as we tried to find a version of their result that was easier to interpret meaningfully in genetics.

Finally, our generalization of the result to the diploid case is interesting from the perspective of both transient and long-term cumulative diploid-mixability, since the diversity can remain even asymptotically in this case. As such, the diploid-mixability will not be the same as fitness asymptotically, and our Theorem thus tells us something that is informative even asymptotically in the diploid case. Even something like Fisher's theorem of natural selection doesn't tell us anything about asymptotic cumulative mixability in the diploid case, so even in this case the connection between MWU and population genetics gives us a new kind of tool to analyze the diploid case.

Chapter 5

Evolution of Animal Personality

We now introduce a model of Animal Personality, and how it could evolve and emerge as a result of MWU applied to various complex personalities. The writing was done primarily by myself and Nina Fefferman, and the so too was the modeling work. The analysis and proofs are my own. How do complexes of such traits evolve? The experts are personalities, modeled as multi-layer neural networks. l_i is the expected opportunity cost in an environment, assuming a simple “personality” is used to predict a noisy, non-stationary real-valued signal. The animal uses MWUA.

Introduction

Over the past decade, many studies have examined the concept of animal personality traits (PTs), defined as tendencies in behaviors (e.g. boldness, aggressiveness, sociability, etc.) that are consistent across environments or contexts, but that vary from individual to individual. These studies have ranged from empirical investigations into distributions of variation in PTs among individuals in a single population in response to particular challenges/scenarios [114] to exploration of the likely ecological contexts that might favor coherent suites of positively correlated PTs (sometimes termed behavioral syndromes [12, 63]). Conceptual models have been built to explore how individual experiences / life histories might be expected to shape individual PTs [39, 128], and how to quantify statistical measures that could test resulting conceptual hypotheses from observations of behaviors/PTs [128]. There have also been efforts to both quantify the fitness benefits of particular PTs [22, 148] and the benefit to individuals participating in populations with a diversity of personalities represented [16, 149]. Often the positive

correlations between different personality traits contribute to the evolution of seemingly maladaptive behaviors, for instance pre-copulatory sexual cannibalism in fishing spiders [66]. Also, under predation sticklebacks evolve positive correlations between boldness and aggressiveness [12], but not under other conditions. What conditions regulate these complex interactions between PT's? Moreover, there is robust evidence that the positive correlations between PT's even affect the nature of interactions between species and other ecological factors [127], suggesting the necessity for a theory that ties together ecological effects and the evolution of PT's. In addition, there has been interesting work showing that the mixture of different PT's within the same species can profoundly affect fitness outcomes, by introducing the possibility of “keystone individuals” who profoundly change group dynamics [129]. Finally, there has been ample evidence that animal personalities affect such factors as feeding and mating behaviors [98, 148] in complex ways that are correlated with other factors, so fitness seems to be affected in a less straight-forward way than the simplest PT theory would suggest.

While each of these studies have explored a critical piece of the puzzle of how animals might develop, and then benefit from, their own individual set of PTs, they fall mostly into three broad categories:

1. Genetically Determined Behaviors — this is the most traditional set of approaches, in which PTs are encoded in an individual's genes. Selective maintenance of each PT is then considered based on the fitness the trait confers in the (potentially fluctuating) environment, and relative to the fitness of other PT genes represented in the population. The emergence/maintenance of diversity among PTs can then be studied using standard techniques from evolutionary game theory [126, 33].
2. Epigenetics — this is the most mechanistic of the approaches, focusing primarily on how individual life history may critically affect the expression of genetically determined behavioral traits [117]. In this way, individuals who may all share a common set of genes will express different consistent PTs due to the impact of their experiences on the regulation of gene methylation/expression.
3. Life history-based Learning — this is the most nuanced of the approaches, in which

individuals experiment with particular choices, the results of which either reinforce or diminish their likelihood of making similar choices in the future [150, 128]. Taken in this way, PTs become a mechanism for behavioral plasticity, allowing individuals to converge on responses that are beneficial in their environment, based on their individual experiences, and then behave consistently thereafter. Each early choice suggests inferences about their environment, the individual's relative ability within that environment, and the individual's fitness benefits from making particular choices within that context. Diversity of expressed personalities can then be explained in a number of different ways: by stochastic sampling of a fluctuating environment, by differences in individual capability (note: many studies have assumed consistent differences in productivity across individuals in the same environment; e.g. [16]), and the order of early behavioral experiments.

Of these three, only the last incorporates a mechanism for active choices, made by individuals to increase fitness. However, even in discussions of life-history based learning, the elements of learning have only been applied to developing PTs reflecting particular behavioral actions (e.g. learning to 'be bold') and not to underlying algorithms by which individuals might select among possible actions. Without requiring any assumptions about animal cognition, there are likely to be underlying algorithms governing decision-making across contexts. These algorithms can range from those as simple as 'evaluate some proximate signal and behave deterministically based on the result', to those that rely on actions of conspecifics, such as 'mimic the currently observable actions of others' (typically premised on the idea that performers of unsuccessful actions are less available to observe due to active selection), to those that involve simple learning in a juvenile state, such as 'mimic actions remembered/learned from parents', all the way to active, iterative algorithms that involve learning, such as 'try something at random to gain life history experience for a while and then refer to life history experience thereafter'. (Of course, animals who differ in the underlying algorithm by which they make decisions could be indistinguishable from animals who learn to take particular actions based on life-history learning feedback unless experiments are specifically targeted at teasing the two apart.)

By shifting the focus in this way, away from PTs or behavioral syndromes, to general algorithms by which individuals make many types of choices, we can now begin to discuss personality as the way in which individuals interact with their environment and arrive at behavioral decisions. This also allows us to abstract our discussion of the evolution of the capacity for animal personality away from the fitness conferred by particular PTs themselves and ask how can selection have favored the emergence of the trait “having a personality?” To the best of our knowledge, no study has yet provided a unifying theoretical model exploring the evolutionary selective pressures that would lead to the capability of animals to form more nuanced personalities than those simply described by “responsive to their environment” vs “unresponsive to their environment” [149] . We here borrow from an area of Learning Theory in Computer Science called Multiplicative Weight Updates Algorithm (MWUA) and show how the combination and interpretation of existing results provide both a natural way to discuss personality itself and to evaluate the relative fitness of formation of personality. In this way, we generalize the insight from the literature on life-history based learning favoring the evolution of particular PTs to the evolution of personality itself.

Results

Model

Many of the questions of interest in the evolution of Animal personality are complex and multi-faceted, as presented above. Many of the complexities can be viewed as different facets of one model based on an expanded account of adaptive personalities. In particular, we consider that correlations between traits in evolutionary time and between individual (behavioral) variance and average environmental variance, could come about as a result of a learning process during a (not-necessarily pre-reproductive) “juvenile state.” Such a model allows us to use the methods of learning theory to make testable predictions about adaptive personalities. According to the learning model, during the juvenile state, the organism will be able to try different personalities, and evaluate them based on their opportunity cost to the organism in a variable environment (defined as

sub-optimality in relative fitness compared to the optimal fitness). Then the animal tries out different personalities probabilistically, based on their estimated opportunity cost. (We assume that the animal is trying to estimate the opportunity cost of the different personalities based on tracking ecologically-relevant signals in the environment.) Of course, an animal's estimate of the opportunity cost of a given personality can only be a noisy estimate, as a result of the following factors:

1. Environmental fluctuations
2. Strength of the signal to noise ratio for the ecologically-relevant signals
3. The ability of the animal to discern appropriate metrics for the signals
4. The efficiency of resource utilization or responsiveness, which could vary between different animals

The (estimated) opportunity cost of the personality, as experienced during the juvenile state, then proportionately decreases the probability of choosing that personality. Based on this model, we can ask how long should the juvenile state last for any given species? In the animal personality literature, it is mentioned that personality is impacted by estimated opportunity cost for a fixed period in which environmental challenges occur [135, 37]. We generalize that insight to characterize how long the juvenile state should last in order to identify a personality that balances successfully the tradeoff between the relative cost of continuing to search and the marginal returns on the fitness.

Using such a theory, we should be able to make testable predictions across particular populations/taxa about how long the juvenile state should be based on statistical characterizations of fluctuations of their environment/ecosystem. Of course, to test predictions of any model about the length of juvenile state, there is a trade-off between the insight gained and the number of experiments necessary to verify the prediction. Thus the diminishing returns of such experiments, as far as testing these predictions, must be understood before studying models of the juvenile state, including ours.

The Learning dynamics

Assume there exists a suite of potential algorithms for decision-making (personalities) from which an individual can choose probabilistically, inversely proportionate to estimated opportunity cost. For each personality, we can associate an expected opportunity cost for performance in the environment in which the individual lives, incorporating noise in the estimated opportunity cost to reflect environmental fluctuations, the signal/noise ratio of the ecologically-relevant signals, and between-animal variability. Each personality has an estimated opportunity cost, but so does the individual, based on the average opportunity cost of the personalities chosen. We model the ecologically-relevant signal in the environment as a stationary Markov Chain with known mixing coefficients β_a for all $a > 0$. The model then assumes there is a supervised learning or regression procedure for each personality, with each personality predicting the next value of the ecologically-relevant signal from the past history. Assume that each algorithm for behavioral choice has an inherent associated memory length, d . We assume that the algorithms used are different kinds of multi-layer feedforward neural networks [4] with sigmoid transfer functions with l layers, and each unit i has d_{i-1} inputs. Consider these models of personalities to be a generalization of quantitative phenotypic traits which also allows for other quantitative models in general. Let $N = \sum_{i=1}^l \sum_{j=1}^{d_{i-1}} d_{i-1} + 1$ be the number of parameters of the neural network.

It is assumed the “juvenile state” is of length n . Let μ and a be constants such that $2\mu a + d \leq n$. Let B be the upper-bound on the average opportunity cost for any personality with respect to norm q given all likely ecologically-relevant signals in the environment. There is a training period during the juvenile state, during which the animal uses the MWUA to adjust the probability of using a mechanism. The estimate of opportunity cost for personality m based on the experienced states we call \hat{l}_m . Let us define now ϵ , the total “credence” in the current response of a particular personality, describing to what degree the model can be shifted in the face of new local information from the environment:

$$\epsilon = \frac{2B\mu\tau(q)}{\sqrt{\mu}} \sqrt{O(N \log N) \log(1 + 2\mu/O(N \log N)) - \log(\delta'/8)}$$

where $\delta' = \delta - 2(\mu - 1)\beta_{a-d}$, and $\tau(q)$ is a monotonically increasing function of q . The true opportunity cost for the regression we denote $R_n(m)$, and the estimated average opportunity cost for m is $\hat{R}_n(m)$ (averaged after n data points). Now the following holds with probability at least $1 - \delta$ for all $\delta > 2(\mu - 1)\beta_{a-d}$ (see the SI for proof):

$$\hat{R}_n(m) = (1 - \epsilon)_+ \theta(R_n(f))$$

Then it holds that using the MWUA (with temperature η) to determine the mechanism usage probability (from a set of mechanisms M), we obtain, with the same probability as above (see the SI for proof):

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} \leq \sum_t l_m^{(t)} (1 - \epsilon)_+ + \eta \sum_t |l_m^{(t)} (1 + \epsilon)_+| + \log |M| / \eta$$

where m is the personality with the least opportunity cost. So by using the MWUA the animal personality is close to having minimal opportunity cost, but if ϵ (the credence) is large, then the animal's opportunity cost is large. As the animal is using the MWUA, the opportunity cost for choosing the best personality type using the MWUA is therefore the relative fitness of the organism.

The features on which the animal's opportunity costs rely include: how complex and multi-layered the response mechanism for the personality is (N); how hard it is to predict the environment given the memory constraints (β_{a-d}); the number of personalities ($|M|$); the responsiveness of the animal to evidence that a personality is not adaptive ($1/\eta$); and the the maximum possible opportunity cost given the most likely states of the environment (B).

Discussion / Conclusion

The MWUA models allow us to unify some concerns in the Animal personality literature, address some of the unresolved problems, and further allow us to address new,

previously inaccessible aspects.

A recent evolutionary model [149] for the emergence of responsive personalities gives circumstances under which personalities might respond to the environment. In our framework, the general issue is treated by considering and characterizing the animal’s credence η in its current reaction strategy. The responsivity of a personality is therefore modeled smoothly by its credence η , and (for instance) one could also add personalities that are “unresponsive”, for which there is no adjustment to the environment. In this case, one could show using the relative fitness of the animal personality for our model that any responsive personality would only be chosen if it does better than the unresponsive personality (see Remark 1 in the SI). However, we show that there are many other factors besides the benefits of responding that could affect the use of responsive personalities, including how multi-layered and complex the personality is.

Broadly speaking the feedback-based personality theory is also treated in this model, as state-behavior feedback is here treated in the multi-layered models used for each personality (the internal nodes in the neural network can be viewed as internal states) and the behavior ends up predicting the environmental cue or not. Some questions posed in investigations into that theory can also be partially characterized by our framework. For instance, what is the timescale over which the feedback loops act? We characterize this using the estimate of how long the juvenile state should last to attain a desired balance between search time and marginal returns on fitness (see Remark 2 in the SI).

More generally of course the model characterizes how many personalities would emerge due to the state-behavior feedback (another highlighted issue in the literature), as the opportunity cost of the animal decreases, in our model the number of possible personalities chosen increases (see Inequality 1 in the SI for the tradeoff between number of personalities and the opportunity cost of the animal personality, one is on the RHS of the inequality, the other is on the LHS).

Finally, we can characterize when positive feedback generates among-individual differences in fitness, as opposed to alternative state-behavior combinations with equal fitness. To see this, consider that we can give circumstances under which the number of personalities at the end of the juvenile state is larger than one (using the opportunity

cost estimate again) and also check how this is affected by having state-behavior combinations of the same fitness (again see Remark 2 in the SI for the right estimate of the juvenile state length). In particular, for equally-fit state-behavior combinations, one would expect that the state-behavior combinations, since they have lower complexity N , would be more favored when the environment is quickly changing and not much can be done to predict the changes in the environment (note how the support of the distribution over p_i is affected by the length of the juvenile state under the stated condition). Thus the reduction of loss associated with the increase in complexity N is not very big. In addition, there is a tradeoff between generation length and juvenile state length (the volatility of environmental fluctuation and the memory length / sampling of personalities). For instance, imagine that the environment is not fluctuating, then the relative cost of choosing the adaptive personality with a juvenile state may be too forbidding.

Any discussion of the evolution of animal personality must simultaneously address two issues. The first is: ‘If particular PTs and/or personalities are evolutionarily beneficial, why would selection not favor their uniform adoption?’ Of course, the most natural answer to this is that environmental and ecological conditions fluctuate rapidly enough that the ability of individuals to shape their behaviors in response to current conditions (rather than those that determined the fitness of their parents) confers fitness benefits. The second issue is therefore naturally: ‘If the environment is so highly variable that particular PTs are not consistently beneficial, how can the existence of PTs/personality increase fitness rather than the consistency in behavioral choices itself compromising the ability of the individual to react to the environmental conditions?’

The MWUA models allow us to elegantly and rigorously explore exactly those evolutionary conditions in which these tradeoffs are most successfully balanced by employing personality as the mechanism for behavioral decision making in response to shifting environments. We are able to consider selection acting on the lengths of juvenile states, make predictions about the relative numbers of distinct personality types we should observe in species depending on the magnitude of impact to fitness expected from adopting aligned vs. mismatched personalities in the environment, and explore the impact

of duration of memory in life-history based learning. Moreover, the MWUA is itself a mathematical characterization of a process of learning and selection, meaning that the behavior of such a system is not just a quantitative description, but demonstrates how evolution could lead to the emergence of personality proper.

5.1 Supplemental Information

We think of the organism choosing between different personalities according to the Multiplicative Weight Updates Algorithm (MWUA). Theorem 2.5 in Arora et al. states that, under the MWUA, for any agent i :

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} \leq \sum_{t=1}^T l_i^{(t)} - \eta \sum_{t=1}^T |l_i^{(t)}| - \frac{\ln n}{\eta}.$$

For large T , the last term becomes negligible, and for small η , the second term is small. Now, since this inequality holds for any expert i , it holds specifically also for the best expert in retrospect, and thus the performance of the MWUA is close to that of the best expert in retrospect.

We model each expert i as an animal personality, and the losses \hat{l}_i are “performance” or predicted loss of the animal personality, with the following setup. The loss is the fluctuation in the environment that influences the fitness payoff from the behavior of the animal personality in response to a changing environment:

The animal is trying to predict some environmental change with the environment sampled from a stationary markov chain with mixing coefficient β_a for all $a > 0$. During the juvenile state the organism estimates which distribution over personalities to use, with each personality being a parametrized multi-layer neural network. The average empirical loss $\sum_t \hat{l}_i^{(t)}$ is the average loss of personality i during the juvenile period. The true loss for a personality is approximated by the average empirical loss associated with that particular set of parameters for the neural network model given the environmental interactions during the juvenile period. The loss incurred by the organism at time t is $l^{(t)}$. The organism is assumed to use MWUA to select the best personality based on the average empirical loss.

The critical parameters in the environment that affect fitness are being tracked by the organism, and are assumed to be real-valued. Assume that each mechanism has memory length d , pooling observations for that period of time for each data-point. We assume that the personalities used are different kinds of multi-layer feedforward neural networks having sigmoid transfer functions with l layers, and each unit i having d_{i-1} inputs. Let $N = \sum_{i=1}^l \sum_{j=1}^{d_{i-1}} d_{i-1} + 1$ be the number of parameters of each personality's neural network.

It is assumed the "learning period" is of length n . Let μ and a be constants such that $2\mu a + d \leq n$. Let B be the upper-bound on the possible error for any personality with respect to norm q . Let

$$\epsilon = \frac{2B\mu\tau(q)}{\sqrt{\mu}} \sqrt{O(N \log N) \log(1 + 2\mu/O(N \log N)) - \log(\delta'/8)}$$

where $\delta' = \delta - 2(\mu - 1)\beta_{a-d}$, $\tau(q) = \sqrt[q]{\frac{1}{2} \left(\frac{q-1}{q-2}\right)^{q-1}}$ is a monotonically increasing function of q . We assume the following upper-bound on the possible loss for the environmental model holds. For all $f \in F$ and some $q > 2$:

$$1 \leq \frac{(E_P[(l(f(X), Y)^q)])^{1/q}}{R_n(f)} < B$$

Such a bound is more general than a loss that is bounded, and includes even some heavy-tailed distributions for the loss. The true error for the regression we denote $R_n(i)$, and the estimated average error for i is $\hat{R}_n(i)$ (averaged after n data points). Now the following holds with probability at least $1 - \delta$ for all $\delta > 2(\mu - 1)\beta_{a-d}$ (by Corollary 6.5 in [93], and that $VCD(F) = O(N \log N)$ [4], and real valued predictions have an extra degree of freedom):

$$\hat{R}_n(i) = (1 - \epsilon)_+ \theta(R_n(i))$$

which is derived as follows. Corollary 6.5 in [93] states that with probability at least $1 - \delta$ for all $\delta > 2(\mu - 1)\beta_{a-d}$

$$R_n(i) \leq \hat{R}_n(i)/(1 - \epsilon)_+$$

by the def. of R_n , $\hat{R}_n(i) \leq R_n(i)$:

$$\hat{R}_n(i)/(1 - \epsilon)_+ \leq R_n(i)/(1 - \epsilon)_+ \leq R_n(i)$$

Then it holds that using Multiplicative weight updates (with temperature η) to determine the personality usage probability (from a set of personalities M), we obtain, with the same probability as above (by Theorem 2.5):

$$\sum_{t=1}^T l^{(t)} \cdot p^{(t)} - \left(\sum_t l_m^{(t)}(1 - \epsilon)_+ + \eta \sum_t |l_m^{(t)}(1 + \epsilon)_+| \right) \leq \log |M|/\eta \quad (5.1)$$

where m is the optimal personality. The negative LHS gives the relative fitness of that organism. We assume that in the case of zero-loss, the fitness is equal to s , and the relative fitness is the deviation of our current genotype's fitness from that of the (not-necessarily zero-loss) best possible genotype. The following factors increase ϵ and thus decrease the relative fitness: the complexity/nonlinearity of the personalities N , the rate of decay of correlations (mixing coefficient) β_{a-d} at the time-scale at which the optimal animal personality experiences and recalls stimuli $a - d$, the number of mechanisms $|M|$, the upper-bound on average error for any personality B , and the degree of “temerity” $1/\eta$ the animal has for choosing the best personality.

Remark 10. *As Inequality 5.1 for the relative fitness of the personality shows, the average loss of the animal personality approaches that of the best personality.*

Remark 11. *To estimate the length T of the juvenile state necessary for the animal personality to achieve a fixed level of cumulative relative fitness, one can derive using the above:*

$$\frac{1}{T} \sum_{t=1}^T l^{(t)} \cdot p^{(t)} - \frac{1}{T} \left(\sum_t l_m^{(t)}(1 - \epsilon)_+ + \eta \sum_t |l_m^{(t)}(1 + \epsilon)_+| \right) \leq \frac{1}{T} \log |M|/\eta$$

and set T as a function of $|M|, \eta$, and ϵ to achieve an appropriate desired upper-bound for the time-normalized regret on the quantity

$$\frac{1}{T} \left(\sum_{t=1}^T l^{(t)} \cdot p^{(t)} - \left(\sum_t l_m^{(t)} + \eta \sum_t |l_m^{(t)}| \right) \right)$$

and by so doing, one can get the final upper-bound on the cumulative relative fitness by multiplying both sides by T and adding $\left(\sum_t l_m^{(t)} + \eta \sum_t |l_m^{(t)}| \right)$ to both sides.

For the length of the juvenile state, rather than using a relative fitness optimality criterion, we look at marginal returns on potential search times. In particular, we consider the marginal benefit of continued search time (as opposed to previous time-points) instead of the relative fitness. Using marginal fitness as the metric of optimal return. To be clear, the juvenile state length is not the one which gives the personality with minimal opportunity cost, just one which appropriately balances the tradeoff between the relative cost of sampling vs. continuing. Using the above inequalities one can give estimates of how long the juvenile state needs to be until there are diminishing marginal benefits of continued search time (e.g. when the cumulative benefit so far is balanced against the cost of continuing).

Chapter 6

Evolution of Tool Innovation

We now describe the model for tool innovation and its evolution using the MWU method. The principal writing in this chapter was done by myself with revisions proposed by Nina Fefferman. The analysis and tools were my own, and the modeling work was jointly done with Nina. Crow tool-innovation is mysterious because New Caledonian crows come up with completely novel tools. The source of novelty we propose is an infinite set of possible tools from which the animal samples, a set which is built up during by playing with similar materials (wood, leaves, etc.) early in life. Every round of experimentation with materials gives a new sample from this set, which then is put in the running as a new “expert” in the MWU. The experts are drawn from a distribution D over the real line with infinite support. The experts are hypotheses h about interactive properties s of the environment (Example: the bucket is hook-able). $-l_i$ is the quality of a hypothesis, modeled as a univariate polynomial over $h - s$. The animal uses MWUA.

6.1 Introduction

Over the past decade and a half, many studies have found evidence of tool-innovation by animals. Studies in the field have included empirical evidence of New Caledonian (NC) crows’ and rooks’ abilities to create novel tools for extraction of food [146, 15]. There is even evidence that a crow is able to replicate an old tool in new materials [147]. American crows are also able to use various types of calls as tools to differentiate between different behavioral contexts, sexes, and many other factors [90, 151]. Surprising findings like these on crow tool-use shed new light on previous work showing the use of tree branches as tools by chimpanzees to extract termites [44]. Conceptual

models exploring the potential mechanisms of evolution of tool use have been proposed, focusing mostly on advanced precursor behavior [2], social evolution [61, 17, 62], and combination play/learning [73, 72].

Precursor behavior models tend to focus on behaviors, like food caching, that are both simple and have direct positive fitness. The precursor behaviors then lead to availability of materials and behaviors that set the stage for easy accidental discovery of tool-use. For example, gathering twigs to build a nest uses the same behaviors that then lead to manipulating twigs for other purposes: e.g. tool-use. Rooks' tool-use abilities are explained well by these food caching behaviors.

Social evolution accounts of the evolution of tool-use suggest that behavioral innovations may happen in many different individuals and then, due to imitation or social diffusion, they spread and are combined. Studies have shown that Whale songs are innovated in this way, with evidence of both social learning/transmission between populations, and also the gradual change in songs in populations [34, 52, 105, 112].

Evidence of social evolution enabling tool use also includes army ants' bridge-building process and Fire ants' raft construction procedure. Fire ant groups make rafts to cross bodies of water, made of the ants themselves [49]. Army ants build bridges made of the ants themselves to cross gaps in the forest floor [118]. Bridge-building and raft-building with ants involves altruism and behaviors that are good for the colony as a whole (common processes in social evolution). The fitness benefit of such structures for those in the colony is due to the bridges/rafts being a structure that allows for crossing obstacles to gather food. Social evolution in this case enables the formation of these bridges and rafts, which are tools that enable crossing bodies of water. Among the Maasai tribe, there is evidence that certain kinds of social risk-pooling behavior requesting need-based transfers based on cultural practices increases survival rate [57]. Moreover, a model inspired by need-based transfers shows that such social schemes can introduce great inequality and thus require strong forms of altruism to persist [58].

Finally, theories that emphasize tool use as combination play with learning generally assume that some propensity for playing with materials and combining them is paired with an ability to learn that a novel combination is useful. The kind of combination

play considered includes things like attempts to manipulate materials with the beak via variations in beak movement/position (in crows; [72]). The kind of learning required for such accounts of tool-innovation yield more generalization than simple associative learning. In particular, there is evidence that generalization from one kind of material to another of the same tool happens very rapidly for the same NC crow with variations in beak movement/position [147]. Combination play is thought to give “experiments” that allow spontaneous invention of tools by NC crows when combined with learning. NC crows form hypotheses about the pliancy properties of the metal, and then combine those with hypotheses about the appropriate position of the tool relative to the openings of the object to be moved. All of these things are done in order to receive food, but are far beyond reinforcement learning [147]. The crow in the generalization experiments was able to generalize to the new materials by using an entirely novel bending movement [147]. The studies of the NC crows explicitly suggested the need for a model that would bring about these novel (for the species and organism’s own experience) behaviors [147]. Persistence over many trials of combination play are necessary for the acquisition of tool use, and many other corvid species do not seem to exclusively use tools except as a kind of temporary foray from food-caching (such as the rooks; [15, 73]). Clearly, therefore, the ability to form hypotheses isn’t itself sufficient to enable the evolution of tool use. Experiment time is necessary during a learning period.

There are a number of elements in the NC crow studies that present additional complexities to be resolved by any proposed model. As interpreted by the authors of the NC crow studies, the learning mechanism used by the NC crows still recruits trial-and-error learning and doesn’t involve perfect insight, as when the correct kind of tool is made, it is sometimes still used improperly. We suggest that the right tools can be made and used improperly for purely exploratory purposes. More specifically, consider that “inappropriate use” of pre-cursor behavior or “correct” tools for the task could be of benefit for future tasks, signaling, or social evolution. For the sake of future tasks, one could engage in improper use of tools to gather information about the materials available in the environment to “practice” for unforeseen contingencies. A use of improper tool use for signaling could give novel social cues. Finally, for social

evolution accounts of tool-innovation, improper tool-use generally (not in the lab) can be understood as an action supporting social group cohesion, novel group participation, or identification of an individual animal to others.

Tool use in the NC crow experiments happens gradually over many experiments [146, 147], and doesn't always develop or remain prominent for other species. For example, neither the rooks [15] nor chimpanzees [113] learn the proper use of certain tools, when NC crows do. For chimpanzees, this is thought to be because they have more general capacities (general aptitude), predisposing them to generate less relevant hypotheses and fewer experiments [147].

We add that, due to social imitation in chimpanzees, less individual experimentation would be expected, but would be done in parallel among different individuals in the same group, thus combining social evolution and combinatorial play/learning accounts. Social learning and mimicry, for instance, can reduce the number of experimental trials necessary for a population of the same species to develop a tool by allowing for collaboration and synthesis from a set of disparate tools.

Given the complexities of the different conceptual models, and the specific call in the existing literature for a more flexible and comprehensive learning model, we provide a new framework that allows us to study all of these factors' contributions to tool innovation and how they could evolve. The starting point is, indeed, the view that is held in common between two of the theories (social evolution and combination play/learning): learning is important, and so is generation of new behaviors. But equally important is the view promoted by the pre-cursor and combination play accounts: that opportunities for experimentation with materials are crucial. We propose that animals evolve tool innovation by evolving the ability to generate novel hypotheses about the hidden properties of objects in their environment, and then learning which hypotheses are correct. Crucially, we use a novel strain of learning theory (Multiplicative Weight Updates on infinite hypotheses) that allows for completely novel hypotheses, and discuss how all three previously proposed conceptual models are compatible with different regimes of this new model.

6.2 Model

Note that, though here described in application to a single individual, the model can be applied just as readily to any con-specific group of animals. First we will describe our model of tool-innovation and the assumptions made by us. Then we will lay it out more formally.

6.2.1 Description and Assumptions

To model the evolution of tool-innovation, we build on the insights of the combination play literature; in keeping with that literature, we assume that the stimuli available to the animal are relatively stable and that the animal has the ability to engage in extensive combination play, attempting many experiments with objects in the environment. Further, the environment has some hidden property that can be discovered by the animal, after which it can be exploited via tool-innovation. In gaining experience through normal interactions with, and exploration of, its environment, the animal forms a hypothesis about the exploitable property. During each round of combination play, the animal draws a new hypothesis from a probability distribution over possible hypotheses, forming a trial hypothesis (at the disposal of the animal during the task). After the trial hypothesis is incorporated into the suite of available hypotheses, all of the hypotheses in the suite are compared to the stimuli, and the best hypothesis is given more credence. After many rounds of play, the animal is assumed to act in accordance with the most-believed hypothesis. The hypotheses about exploitable properties in the environment appropriate to completing the given task can be used by the animal to select fitting actions for fashioning the raw materials into the proper tool (to exploit the property identified by the hypothesis). As an example, consider a small lid with a handle covering a bowl (containing food) placed next to pliable metal strips. The exploitable properties of this food-gathering task include the pliability of the metal in the strips and the fact that the lid can be lifted by a hook-like instrument (it is hook-able). The appropriate actions for such exploitable properties are animal-dependent, but generally are those that bend the pliable metal strips into a kind of hook, and put

the hook in the handle, using the other end to lift the hook and thus the lid, giving access to the contents of the bowl. It is assumed that, before the tool-innovation task, animals are able to engage in combination play during a long “juvenile period,” during which the animal builds up a probability distribution over all possible hidden properties of the stimuli (note that reality suggests the need for unexplored possibilities). We further assume that the hypothesis distribution involves similar tasks to those faced in the tool-making scenarios. Furthermore, assume there is a diverse assortment of raw material available which can be combined to make tools. Critically, there must be some fitness benefit expected from successful tool-innovation. Crucially, there are assumed to be an infinite number of trial hypotheses at the animal’s disposal, which is a source of truly novel hypotheses.

6.2.2 Formal description

The animal keeps a set H of hypotheses in memory, and then gives credence to them based on a probability distribution, choosing h with probability p_h , where $h \in H$. Let D be the probability distribution over hypotheses over a set of possible hidden properties, S , which has infinite size. The distribution D is therefore the probability of the animal guessing a hypothesis $h' \in S$ when first encountering the task. Then, when an experiment of combination play is initiated in a task, a novel hypothesis s is sampled from the distribution D , and added to H as an option with a small probability p_h of being chosen (and then with p_h appropriately re-normalized). During each experiment, the animal is assumed to sample a large number of hypotheses according to p_h (large enough for all hypotheses $h \in H$ to be sampled). Then for each hypothesis h sampled, the high quality $f(h, s)$ of h as an approximation to the hidden property $s \in S$ is assessed based on exploratory actions involving the optimal use of raw materials given that h is the true hidden property explaining the stimuli (roughly speaking, this estimation of quality is based on the use of trial-and-error with makeshift tools constructed). After an experiment is over, then the animal is able to get an estimate of the quality of each hypothesis. Using the estimated quality for each hypothesis $h \in H$, the Multiplicative Weight Updates are used to update p_h . After an experiment, the hypothesis distribution

D can also be updated, but in most of our use of the model we assume it doesn't change (such is the case unless noted otherwise). The quality function $f(h, s)$ is assumed to be a relatively bumpy function (a polynomial), with peaks and valleys, but still smoothly-varying. The main factors in determining how many experiments are necessary for an animal to find a good-enough hypothesis are (1) how many experiments it takes to sample a good-enough hypothesis from D , and (2) granted that (1) happens, how many experiments would it take to identify the good-enough hypothesis as being chosen with high probability p_h assuming that the MWU are used to update p_h after each experiment? The complexity of $f(h, s)$ (see the SI for the formal definition) is noted as $deg(f)$, and gives an estimate of how "rough" the fitness landscape over hypotheses is. If the distribution D over hypotheses has higher probability in some subset $\Omega \subseteq S$ more than elements outside of S (far-from-uniform), then assuming there is some $h' \in \Omega$ that is good-enough in quality compared to the maximum value of $f_{max} = f(h, s)$ over all h , the task takes vastly fewer experiments. By this logic, when the animal is apt to have very well-informed selections for trial hypotheses, then fewer experiments are needed to get a good hypothesis and thus build appropriate tools for the task. For the model, if there are only a few or moderately many likely (according to distribution D) trial hypotheses that are good-enough in quality compared to f_{max} , then the number of experiments necessary until a good-enough tool can be reliably constructed is very large. Alternately, one could have a quality function f that has high or medium complexity $deg(f)$, and the same holds. To combine the two we could refer to the effective complexity $deg(\hat{f})$ as the complexity of the most likely hypotheses according to D (See the SI for the definition). Naturally, the number of experiments available for an animal trying to figure out how to construct an adequate tool for the task is functionally limited. As a result for many tasks, in species with very bad individual guesses (the distributions D which require many experiments to find a good-enough tool), the animal could simply fail to construct an optimal tool, or even succeed in constructing an optimal tool, but eventually end up using some other, less effective tool instead due to the difficulty of the task or the lack of experimentation time (see SI Remark 1). Suppose that the distribution D_T changes throughout task T , ending

up with D'_T at the end of the task. If some set of hypotheses that are good-enough for T become too likely as a trial hypothesis from D'_T as a result of the change, one would expect that D'_T would be useless for a task completely different from T . Then of course, in varying environments, where the types of materials available or the task itself are changing, it makes perfect sense to have a trial-hypothesis distribution D'_T that doesn't incorporate too much information gained from any one task T .

6.3 Discussion

Given the complexities of outcomes for studies of animal tool-innovation, and the variety of models for the same, a unified framework for animal tool-innovation is necessary. Especially considering that there are interesting phenomena that seem to fit somewhere in-between the immediate applications of the models, our model provides a way of combining key insights from past models to capture previously elusive outcomes. Recall for instance the behavior of the rooks and chimpanzees during the tool-innovation tasks. Our model captures desired features of the system, for example that complicated tool-innovation tasks are hard to achieve: Consider an animal trying to solve a task of moderate complexity and moderate effective complexity. Then by the results of the preceding section, the total number of experiments required to consistently construct a good-enough tool for the task is too high, resulting in either a failure to make the right tool at all, or a failure to do so with any reliability. Our model also captures the results from the experimental results outlined in the Introduction. The case of the rooks' regression to caching behavior after finding the correct tool involved much less combination play as compared to NC crows according to Kenward [73]. In the language of our model, it is a case in which the correct hypothesis was guessed, but the number of experiments until it was properly identified as a good-enough hypothesis was too high to be realized by the rooks, who barely did any combination play at all. Similarly, recall that isolated chimpanzees failed to complete the tool-making task completed by the NC crow. If we assume that the chimpanzees have large effective complexity for the task, the number of experiments necessary to get a good-enough hypothesis to be considered is too high to be feasible, thus giving the result that chimpanzees fail

completely to construct the right tool. One important insight that falls naturally from the perspective taken in our model is that, if there are multiple individuals refining their set of hypotheses in parallel, but sharing their experiences via social interactions, they will be able to construct the tool after only a few rounds of experimentation from each individual. This leveraging of communal exploration would yield the highest gain when exploratory hypotheses are minimally overlapping between con-specifics, allowing the fewest number of expected experiments. In this way, social evolution can be considered a critical part of the evolution of tool-discovery/innovation, so long as social interactions enable communication about hypotheses and prior experiments or experiences, even if the task is highly complex and/or the period for experimentation is short. For instance, in the case of the tool-innovation experiments with chimpanzees, this account would predict in a social context with other chimpanzees the quick accomplishment of a tool-innovation task (assuming that they are sufficiently diverse in social standing / personality traits) in contrast with the dismal performance of the isolated individual.

Another critical set of insights that follow from the logic of our model characterize the nature of long-term vs. short-term objectives in tool-making, and efficiency trade-offs in exploration and innovation. The case in which the experiments actually change how often the different hypotheses are guessed allows us to understand some of the seemingly excessive use of trial-and-error with the correct tool that has been observed empirically. Recall that in experiments, NC crows would immediately construct the right tool, but use it wrongly. In fact, our model suggests that this might be exploratory behavior for changing the guessing distribution and doing better on future tasks. As mentioned in the description of the model, if one excludes unfit hypotheses from being guessed too readily, one can't hope to do very well in future tasks using that guessing distribution. There is thus also a trade-off between the number of experiments done in one task and generalization to doing well in other tasks (assuming that there is some narrowing of likely guesses after each experiment). In this sense, optimality is conditional on the expected life history of challenges for which effective tools could be fashioned/utilized. The reason we use "good-enough" hypotheses rather than "close-to-optimal" hypotheses for the goal of the animal is that, due to the trade-off outlined

above, one may stop very far short of constructing the best tool for the task in order to make better guesses in future tasks. Further, one should be guessing hypotheses until the outcome is improved relative to the time/effort invested in the continued experimentation, rather than seeking to find the optimal hypothesis, which could be too costly or lead to poor generalization. When considering searching for either the good-enough vs. the optimal hypothesis, finding the good-enough hypothesis should be understood as finding the optimal hypothesis within a unit of time relative to the best hypothesis. If there is a change in fitness trade-offs for the amount of search time vs. the marginal benefit of achieving a better tool, then this affects the definition of 'good-enough'. As is made clear by our model, selection should favor strategies for innovation that take into account the long-term benefit of leveraging failed exploration from earlier tasks in shortening the need for experimentation for new tasks. Selection should also favor strategies for innovation that take into account the potential for diminishing returns from additional investment in time spent searching for a tool for each task.

6.4 Conclusion

While many of the results of our model are purely intuitive, previously existing models have neither explained nor predicted to the same extent these intuitive and observed outcomes. In addition, we provide a model of tool-innovation that accords with the difficult constraint of providing an account that goes beyond simple reinforcement learning, one of the challenges identified by previous work for modeling NC crow tool-innovation. The first step in advancing our understanding beyond the intuitive outcomes is the creation of just such a model. Our model is consistent with previous, narrower theories, and also with empirical observations, but also provides a broad framework that allows for a deeper discussion of the evolution of tool innovation and its relation to learning.

6.5 Supplementary Information

Recall that in the main text, we introduce a model in which an animal is forming different hypotheses h about a hidden property $s \in S$ of the environment. For simplicity, the

environment is assumed to have only one hidden property. The animal keeps its history of imagined hypotheses H . It is assumed that the animal samples from distribution D a new hypothesis from S and adds it to H . It is assumed that D has infinite support. For instance, one could use a suitable real-valued interval for S , and D could have a sub-interval of S as $\text{supp}(D)$. For what follows, let D be a subset of the real-line which is a subset of $[A, B]$.

The quality of h as compared to s is given by a simple function $f(h, s) = g(h - s)$, with $g : R \rightarrow R$ a polynomial of degree d . Now let the maximal value f_{max} of $f(h, s)$ be equal to $g(0)$.

Definition 12. *Let the degree of quality function $f(h, s)$ be $\text{deg}(f) = \text{deg}(g)$. The complexity of the quality function $f(h, s)$ is $\text{deg}(f)$. The effective complexity $\text{deg}(\hat{f})_\delta$ of the quality function is $\text{deg}(g_D^\delta)$ where $g_D^\delta = \{h(x) = g(x) \text{ for } x | p_D(x) > \delta\}$ where $p_D(x)$ is the density for the distribution D and $h(x)$ is an interpolating polynomial for g at the values with high enough probability mass. For $\delta = 0$, we simply write $\text{deg}(\hat{f})$.*

For complicated distributions D , we can specify δ based on capturing most of the “interesting” probability mass, but for simple distributions like the uniform distributions, we can just use $\delta = 0$.

Now recall that we assume each time a hypothesis is sampled from S , then we model the experiments done by the animal as using a probability distribution over H generated by the Multiplicative Weight Updates Algorithm.

If the quality function, when evaluated on the set H is dissimilar enough to the function on S , then it could be that the expert sensing s could lose against a worse expert. It could be that the true loss function is highly non-linear in potential hypotheses.

In other words, the function being estimated is the true quality function, and each point sampled is another point from a nonlinear function. The point sampled is a hypothesis. One could end up sampling only local maxima. Assuming smoothness in the resulting quality function, if the local maximum is the only thing sampled, and the global maximum is far away from that and close to a low quality hypothesis, then the result would be that the best hypothesis would not be chosen.

Since the quality function is a Lipschitz continuous polynomial of degree d , then we get $d(f(h, s), f(s, s)) < K(d(h - s, 0))$ Where $K = \max_{[a,b]} \|g'(x)\|$ and $d(x, x')$ is the Euclidean distance metric.

So using Markov's polynomial inequality [18], we get:

$$d(f(h, s), f(s, s)) < d^2 \max_D \|g(x)\| (d(h, s))$$

So the higher the degree, the higher the error. We will use this kind of argument in bounding the case of hypotheses sampled from the uniform distribution D over an interval $[a, b]$. Assume that the value of $g(s)$ satisfies $g(s) = y^* E_D[g(x)]$.

Then by the Paley-Zygmund inequality,

$$Pr[g(x) > yE[g(x)]] > (1 - y)^2 \frac{E[g(x)]}{E[g(x)^2]}$$

Then by the first mean value theorem for definite integrals,

$$E[g(x)] = 1/(b - a) \int_a^b g(x) dx = 1/(b - a)[P(b) - P(a)] = 1/(b - a)g(c)(b - a) > d \min(c, c^d) \text{ with } deg(P) = deg(g) + 1 \text{ And } E[g(x)^2] = 1/(b - a) \int_a^b g(x)^2 dx < 1/(b - a)[Q(b) - Q(a)]^2 < 1/(b - a)[K(b - a)^2] < 1/(b - a)[deg(Q)^2 \max_{[a,b]} \|g(x)\| (b - a)] \text{ with } deg(Q) = 2deg(g) + 1$$

Now note that when $Pr(g(x) > aE[g(x)]) \geq \theta$, with probability θ we are able to generate a sample with quality at least $aE[g(x)]$. Then if we sample m points from D , there will be a sample of quality $aE[g(x)]$ with probability $1 - (1 - \theta)^m$. If $m = \frac{\log(1/\epsilon)}{\theta}$ then since $\theta \in (0, 1)$, $m = \frac{\log(1/\epsilon)}{\theta} > \frac{\log(\epsilon)}{\log(1-\theta)}$ and therefore $1 - (1 - \theta)^m > 1 - \epsilon$.

Then with probability at least $1 - \epsilon$, within $m = dB \log(1/\epsilon) / \Omega((b - a)(1 - y)^2)$ samples, we have a sample that has quality at least $aE[g(x)]$. But by assumption a solution with quality at least $yE[g(x)]$ is within $\delta = y^* E[g(x)] - yE[g(x)] = (y^* - y)E[g(x)]$ of $g(s)$. So we have that after $m = \frac{dB \log(1/\epsilon)}{\Omega((b-a)(1-(y^*-w))^2)}$ we have a w -optimal quality sample.

Repeating the analysis for the effective complexity \hat{d} , one obtains after $m = \frac{\hat{d}B \log(1/\epsilon)}{\Omega((b-a)(1-(y^*-w))^2)}$ samples we have a w -optimal quality sample in our pool of hypotheses with probability

at least $1 - \epsilon$.

Even if a good hypothesis is sampled from D , MWUA still takes time to converge. To see this note that after T experiments, the following holds:

Theorem 2.5 in Arora et al. [6] says under the MWUA, for any agent i :

$$\sum_{t=1}^T g^{(t)} \cdot p^{(t)} \geq \sum_{t=1}^T g_i^{(t)} - \eta \sum_{t=1}^T |g_i^{(t)}| - \frac{\ln n}{\eta}. \quad (6.1)$$

For large T , the last term becomes tiny, and for small η , the second term is small. Now, since this inequality holds for any expert i , it holds also for the best expert. Thus the performance of the MWUA is close to that of the best expert in retrospect.

Remark 13. *As the probability p_h of using a bad hypothesis h goes to zero as $T \rightarrow \infty$, one could still have a bad hypothesis in $\text{supp}(p_T)$ for shorter timescales, as can be seen in Inequality 6.1. Noting that the number of experts n is actually a linear function of T , the dynamics will actually never converge to the best hypothesis with probability 1, and will always add new hypotheses to the pool of ones that are evaluated. So in the end the fewer experiments that are necessary for the good hypothesis to be in the pool, the more probability mass will be allocated to trusting the good hypothesis and maintaining use of any tools that are associated.*

Chapter 7

Process Information and Evolution

We now describe work in which Cameron Smith and I analyzed the Rivoire-Leibler model of evolution, and in which MWU is used as a primitive to link it to the universal semantic communication model and population genetics. Thus in this chapter we establish that evolution in a quantifiable sense allows a population to actually “acquire” semantic information about its environment and its energy sources (e.g. that the a number of energy sources therein are useful to the organism). The writing for the chapter was done by both myself and Cameron Smith. The modeling, tools, and analysis were done by both of us as well.

7.1 Introduction

There is no guarantee that any given collection of systems will be capable of productive collective interoperation. In the domain of computing, this is due at least in part to the proliferation of many different man-made systems that have been built by different people at different times as well as the ever-changing standards for these systems resulting from fluctuations in the amount of data used or applications required of them. Indeed, everyday experience is consistent with the failure of any ostensible interoperability guarantee when operating system upgrades lead to system failure, hardware drivers break or old software no longer works according to its specification. Therefore, it is important to understand when interoperability without any fixed standards or agreed-upon protocols is possible and at what cost.

One can view potential *semantic* ambiguity as the crux of the problem of interoperability. This is to ask: How can different systems interpret each other’s actions relative to their own with enough accuracy to reliably succeed in performing their respective

functions but without having the same background protocol, language, or linguistic framework to serve as a necessary precondition.

Information theory is a reasonable place to turn in searching for answers to this kind of question. However, Shannon's information theory does not capture semantics, as pointed out by Shannon himself [124] and by Weaver. In particular, Weaver says (as quoted from [145]):

“The effectiveness of a communications process could be measured by answering any of the following three questions

1. How accurately can the symbols that encode the message be transmitted ('the technical problem')?
2. How precisely do the transmitted symbols convey the desired meaning ('the semantics problem')?
3. How effective is the received message in changing conduct ('the effectiveness problem')?”

Shannon information specifically addresses item one from this list without referring to issues related to items two or three.

Semantic communication is addressed by the theory due to Juba referred to as Universal Semantic Communication (USC) [67]. The conceptual advance put forth in USC is to associate a common goal to the objective of communication, which combines items two and three from Weaver's criteria for effective communication above. Thus the emphasis shifts from information being a thing to pass on, to being a kind of process, in which two agents interact and are only successful if they achieve some goal that has an impact on all agents involved. The common goal may allow for an arbitrary degree of cooperation: the goal may be common while favoring the overall interests of one individual over another.

There is much discussion about information-processing in biology in general and evolutionary theory in particular that requires a theory of semantic information [134, 120]. In biology and evolution, Eigen [43] says the following of Shannon information with respect to issues it does not resolve that require a theory of semantic information:

“Information theory as we understand it today is more a communication theory. It deals with problems of processing information rather than of ‘generating’ information. It requires information to be present ‘ab initio’ in a well defined form; It always requires ‘somebody’ – usually man – to determine what to call ‘information’ and what to call ‘nonsense.’ This complementarity between information and entropy shows clearly the limited application of classical information theory to problems of evolution. It is of little help as long as information has not yet reached its ‘full meaning’, or as long as there are still many choices for generating new information. Here we need a new variable, a ‘value’ parameter, which characterizes the level of evolution.”

In trying to answer this call for an interpretation of information in biological evolution, Rivoire and Leibler [120] show that under the assumptions (1) No information is inherited between generations, (2) any information acquired from the environment is common to all members of the population, and (3) only one type predominates in each environment, the long-term fitness of the population, if increased, implies gains in information, in the classical Shannon information-theoretic sense. Furthermore, in the same work they give a theory of semantic information. The importance of Rivoire and Leibler’s model is its usefulness to those using Information-theoretic models in biology, as it shows in some cases mutual information can be equal to fitness. Juba’s USC and Rivoire-Leibler theory have what appear on the surface to be quite different applications in mind, it is, however, remarkable how similar the two are in conceptualizing information as a process whose end point is a kind of statistical learning. We will show that the two are in fact more than just conceptually related, as there is a particular type of USC goal toward which a population evolving according to the dictates of the Rivoire-Leibler model is at least implicitly directed. Moreover, we should note that we take our example from a very restricted class of USC goals, and thus we show a tight connection between USC and the Rivoire-Leibler model (as opposed to using a very general USC model without restrictions, and using the ample degrees of freedom to embed Rivoire-Leibler evolution).

Juba shows that general USC (which he calls finding the generic universal user) is indeed possible, but only with considerable communication overhead. He gives a few different results for special cases [69, 68]. Here we give a different special case of the general problem (Theorem 5), and show that this problem is actually solved efficiently by a kind of population genetic dynamics (Theorem 13).

The connection between a special class of USC problems and evolution is demonstrated through a particular kind of process-based information in which one agent tries to translate a signal from the other (e.g. words of a language that may be initially unique to the sender), assuming the existence of a common interpretation. Our process information model ends up encapsulating a class of USC problems (Theorem 5), and also captures what the Rivoire-Leibler model states about semantics. We use this connection to show that the Rivoire-Leibler model can thus provide a solution to a USC problem. We also generalize prior work on Multiplicative Weights Updates (MWU) and evolution, in addition showing that MWU is a way of generalizing the Rivoire-Leibler model of evolution, and using that result to show that infinite population selection with no mutation and no recombination conforms to the Rivoire-Leibler model (Theorems 9 and 10). Finally, we show that our generalization of the Rivoire-Leibler model, which uses process information rather than Shannon information, also generalizes their insight that an evolving population is learning information about its environment (Theorem 11).

The proof of Theorem 11 relies on the following insight: one can use the Rivoire-Leibler model to make equality comparisons between the fitnesses of different phenotypes by using the reward function given by the process information model (to construct a suitable Rivoire-Leibler process). Then one can use the comparators as a primitive to run through all possible pairs of phenotypes, and find out which phenotypes are in different equivalence classes. The equivalence classes of fitnesses can then be used to solve the process information problem, by decoding which word pairs mean the same thing. One can then further use this kind of information to implement a protocol for a simple USC problem. The interpretation of Theorem 11 is that the Rivoire-Leibler

model can simulate the acquisition of process information, with the environment having an unknown semantics for which states are compatible with which phenotype, and the population deciphering equivalent compatibility (environmental meaning) of phenotypes.

So, in sum, asexual selection with no mutation or recombination in the weak selection regime (and strategies for the Rivoire-Leibler model in general) can be interpreted as an algorithm for learning which phenotypes have equal fitness in which environment states. That is, we can interpret the information being gained in the Rivoire-Leibler model as being "about" the members of these equivalence classes w.r.t. fitness.

7.2 Process information

Before defining process information, we describe some background concepts. Informally, we assume that there are two agents Alice and Bob who don't share the same language, but who have the same collection of concepts associated with the words comprising their respective languages. The words signify or refer to the concepts (in this model, what words signify is what they mean). Words of Bob and Alice that signify the same concept describe some way of answering "what is it?" questions in the world, which we refer to as *forms*. This model of language is called the triangle of reference [107, 5]. Examples of forms would be things like colors, dogs, or stones. For instance, a form would be the color corresponding to human perception of light with wavelength ≈ 475 nm. Alice speaks English and Bob speaks Spanish, so Alice uses the word "blue" and Bob uses the word "azul." Both have a concept corresponding to human perception of light with wavelength ≈ 475 nm which comes to mind when they hear "blue" and "azul" respectively. The concept of blue then is the concept that Alice and Bob have of the form blue, based on their experiences of elements possessing that color. The concept only corresponds to the form if it is accurate, and the two words for the concept reliably come to mind when the form is shown to Alice and Bob. For simplicity, we further assume that there is at most one form for every pair of words, and at most one concept for every pair of words that correspond to the same concept. The relationships

between words, a form, and a concept can be visualized as:

$$\begin{array}{ccc}
 \textit{Form} & \longrightarrow & \text{Bob's word} \\
 \downarrow & & \downarrow \\
 \text{Alice's word} & \longrightarrow & \textit{Concept}.
 \end{array}
 \tag{7.1}$$

Each of the arrows in Equation 7.1 we understand as a kind of transformation or process. The intuitive colloquial conceptual treatment of this diagram is usefully formalized most generally in the context of category theory. We will, however, use set theory for this section, as it is the simplest way of introducing our model of semantics (and all that is necessary for understanding our the connection between Universal Semantic Communication theory and Evolution). We define the pullback of two functions $f : X \rightarrow Z$ and $g : Y \rightarrow Z$ to be the set $X \times_Z Y = \{(x, y) \in X \times Y \mid f(x) = g(y)\}$. We define the projection maps p_X and p_Y to give copies of the elements in $X \times_Z Y$, only the first indices for p_X and only the second indices for p_Y .

We assume that both Alice and Bob's words are sets that are subsets of a universe C . Alice's words we denote as $A \subseteq C$, and Bob's words we denote as $B \subseteq C$. Alice and Bob are assumed to each have a collection of words associated to the same collection of concepts even though the word that Alice uses for a given concept may be different from the word that Bob uses for that same concept. The elements of this common collection of concepts, the collection to be denoted ψ , are also taken to be elements in C so that $\psi \subseteq C$. Let $Q \subset A \times B$ be a set of (w, w') pairs that refer to or signify the same concept. To model Diagram 7.1 (for Alice and Bob's words and concepts), we assumed that the same concept mapped to multiple words $w \in A$ and $w' \in B$. Furthermore, we model the other content in Diagram 7.1 by saying for all pairs of words $w \in A$ and $w' \in B$ and $(w, w') \in Q$ there exists some concept $c \in \psi$ together with functions $f_A : w \rightarrow c$ and $f_B : w' \rightarrow c$. As there is at most one form for every pair of words (though not every word pair is associated with a form), one way of modeling this is to assume that the form $F(w, w')$ of w, w' is an element of the pullback of f_A and f_B , and since C has pullbacks over C , $F(w, w')$ always exists for all elements in

$(A \cap \text{proj}_1(Q)) \cup (B \cap \text{proj}_2(Q))$, and is unique (see Section 7.8).

The relation between w , w' , $F(w, w')$ and c is given in this diagram:

$$\begin{array}{ccc} F(w, w') & \xrightarrow{p_B} & w' \\ p_A \downarrow & & \downarrow f_B \\ w & \xrightarrow{f_A} & c \end{array}$$

where p_A and p_B are the projection functions from $F(w, w')$ to w and w' respectively. Now we assume that w and w' are publicly available to both Alice and Bob, $F(w, w')$ is private but can be revealed, and f_A , f_B are private to Alice and Bob respectively and cannot be revealed to the other directly. If Alice wants to communicate what w signifies (which is c), therefore she can't show him f_A or c , but instead has to communicate in such a way that Bob can decipher c .

The objective is for Bob to learn all of the associations between the common collection of concepts and Alice's words. One way that Alice and Bob can communicate to achieve this goal is for Bob to use concept-word feedback $r_c(w, w', x)$, which gives real-valued feedback about whether there exists a function whose domain includes $w \in A$ and $w' \in B$ and that has as its codomain $x \in \psi$. $r_c(w, w', x) > 0$ when there exists a function in the set of all functions involving elements in C such that w and w' are in its domain and c is its codomain, and is 0 otherwise. $r_c(w, w', x)$ is equal to 1 when it is positive. In words, $r_c(w, w', x)$ is positive when functions exist from w and w' to x and $x = c$. If $r_c(w, w', x) = 0$ then there does not exist a function between w and x and likewise for w' and x . A more general reward function $r_c(w, w', x)$ will have nonzero values that are real values generally (not just 1), to encode the extent to which w, w' are compatible with or describe c .

Consider a process wherein one system, Bob, is provided with information that may be sufficient to learn all forms for the common terms Q of a pair of systems, Bob and Alice, using $r_c(w, w', x)$. Let $n = |C|$. If we have an algorithm \mathcal{A} that one system, Bob, can use to learn all of the other system's, Alice's, words when given access to r_c in $m(n)$ steps, then we say the first system, Bob, can be $m(n)$ -informed, and \mathcal{A} is

said to be $m(n)$ -informing. We say Bob is *informed* in k steps if there exists an online algorithm \mathcal{A} that Bob can use to learn all of Alice's words in Q when given access to r_c in k steps. The rate of information from initiation until time t is then measured by using the number of words we have properly identified thus far. Whether one system is capable of being informed in this fashion at all and, if so, the number of steps k required to do so together qualitatively and quantitatively characterize what we refer to as *process information*.

More formally:

Definition 14 (Process information). *Bob can be informed, in the sense of process information, in k steps, if there exists an online algorithm with inputs A, B, C and r_c which receives k pairs (w, w') where $w \in A$, $w' \in B$, outputs the set of pullbacks for all pairs $(w, w') \in Q$, using the pairs when combined with r_c , assuming $x \in C$ is the signifier which is given to r_c by the Algorithm \mathcal{A} each round when given (w, w') (whenever \mathcal{A} uses $r_c(w, w', x)$).*

Positive feedback for \mathcal{A} we say is a positive reward $r_c(w, w', x) > 0$. If the algorithm \mathcal{A} doesn't change its state at all when it receives positive feedback, then we say \mathcal{A} is *conservative*. An algorithm \mathcal{A} is conservative, in other words, when it guesses correctly that c is the concept associated with w and w' and doesn't change its state. We note that the notion of a conservative algorithm is vital for our link between a certain class of USC-protocols inspired by machine-learning applications, and was introduced by those authors [69]. We simply apply this notion of conservative algorithms to connect a slight generalization of this class of USC protocols to our notion of process information. The use of such algorithms is so important because the learning-related USC protocols all require this kind of algorithm to be simulated [69]. As a natural generalization of this work, we use a similar simulation technique.

Interestingly, the above semantic communication model can be generalized beyond Set theory to Category theory. The generalization is detailed in Section 7.9. Such a generalization allows us to talk about communication over structured representations that are more complex than simple sets. The background concepts from category theory

we apply in describing the generalization are reviewed in Section 7.8.

7.3 Process information and universal semantic communication

Aside from sharing the conceptual underpinnings of the primacy of process, goal-based notions of communication, and semantics, process information and USC can be formally related. The setting of process information resembles that of online learning, which is not so surprising, since the relation between USC and online learning [69] is very similar to the relation of USC and process information that we demonstrate. What follows is a very brief guide to USC (see [69, 67] for a more complete account), followed by a result that process information as we have defined it corresponds to a certain kind of universal user within the USC framework.

The basic idea behind USC is to have a user and a server accomplish some goal together while interacting with their environment. User, server, and environment each have some internal state, and they are each joined by a (two-way) communications channel that also has a fixed state on each round. User, server, and environment have a strategy that specifies a distribution over new internal states and outgoing messages for the following round, given the entity's current state and incoming messages. Given strategies for each of the entities, the system as a whole is modeled by a discrete-time Markov process with a state space Ω . The resulting stochastic process is given by the infinite sequence of random variables X_1, \dots, X_t, \dots where X_t is the state of the system in round t . See Section 7.7 for more details about the USC framework.

A goal G is a pair (E, R) , where E is a non-deterministic environmental model and R is a referee that classifies whether the goal is achieved. A user strategy $u \in U$ is S -universal with respect to a goal G if, for every server strategy $s \in S$, (u, s) robustly achieves the goal (achieves it for all E and from all initial conditions). Also define for states a size parameter function $sz : \Omega \rightarrow \mathbb{N}$, and a bound $B : U \times \mathbb{N} \rightarrow \mathbb{N}$, taking a user strategy and a target error rate and giving a bound on the number of rounds needed to realize rate ϵ . Using sensing functions (see definition in Section 7.7), USC can guarantee that either $B(u, \epsilon)$ errors will occur or the referee is 0 (safety), or the

referee is 1 (viability) for the achievement of a goal G . More specifically, in USC, one has v -viability for G for a sensing function V if there is a user strategy $u \in U$ such that for all $e \in E$, $\sigma_1 \in \Omega$, V after $v(sz(\sigma_1))$ rounds evaluates to 1 in every subsequent round in the execution of the protocol (where σ_1 is an initial state used by the execution of the protocol) with probability 1. V is s' -safe for G if for all $e \in E$, user strategies $u \in U$, starting states σ_1 , whenever $R'(\sigma_1) = 0$, then $s'(sz(\sigma_1))$ errors will occur, or for some $t \leq s'(sz(\sigma_1))$, V evaluates to 0 in some state X_t . The aim of USC is to give a good class of users, called the generic universal users, that achieve the goal in a robust sense:

Definition 15. *For a class of goals in infinite executions G , a class of user strategies U , we say that $u \in U$ is a B -error (U, s', v) -generic universal user for G if for all $g \in G$, any server s , and any sensing function V that is s' -safe and v -viable with s with respect to U for G when $u \in U$ is provided the verdicts of V as auxiliary input, (u, s) robustly achieves G with $\min_{U_S \in \{U \mid U_S \text{ } v\text{-viable with } s\}} B(U_S, \cdot)$ errors.*

For the result that follows, additional information about the USC model is available in Section 7.7. We prove that process information and a special case of USC are linked by the following result, showing that a solution to a process-information problem gives a generic universal user for a specific class of goals:

Theorem 5. *Let G_C be a class of one-round multi-session goals in which the user's incoming messages on each round are drawn from a set S of sets in C ($S \subset C$), and its outgoing messages are from the set of sets $P \subset C$. Let U be a class of functions $\{f : S \rightarrow P\}$ with a size parameter $n : U \rightarrow N$. Then a conservative $m(n)$ -informing algorithm for U is a m' -error generic universal user over U with 1-viable and 1-safe sensing functions for G_C for error bound $m'(U, n') = m(n(U)) + 1$*

Proof. See Section 7.7 □

Theorem 5 indicates the necessary conditions for a process to support the faithful transmission of semantic information via USC in the sense of process information. In what proceeds we review the connection between online learning and evolution in order to show that the relationship between USC and process information demonstrated in

this section extends to models of evolutionary processes where semantic information is at stake.

The broad outline for the result is that one can use the algorithm that solves process-information in a white-box way, able to set and access the internal order of the pairs of words given to the algorithm, and other internal variables, besides the standard inputs. Then the white-box use of the algorithm allows one to take incoming messages as proposed word-pair inputs for the algorithm, and the outputs of the sensing function as the reward $r_c(w, w', x)$ and so forth.

7.4 Online learning, Multiplicative Weight Updates, and Evolution

Online learning protocols are described as follows [23]. At all time steps $t = 1, 2, \dots, T$:

- Choose action $a_t \in A$
- Simultaneously an adversary (or Nature) selects $z_t \in Z$
- Receive loss $l(a_t, z_t)$
- Observe z_t

The objective of these protocols is to choose a sequence of actions a_t that minimize the cumulative regret:

$$R_T = \sum_{t=1}^T l(a_t, z_t) - \inf_{a \in A} \sum_{t=1}^T l(a, z_t)$$

with $|A| = n$.

Definition 16 (Online Learning). *An online learning protocol achieves no-regret when $\lim_{T \rightarrow \infty} \frac{1}{T} R_T \leq 0$.*

The Multiplicative Weights Updates (MWU) are an online learning protocol that selects actions probabilistically according to the following scheme:

$$p_t(i) = \frac{p_{t-1}(i)(1 - \epsilon l(a_i, z_t))}{\sum_{j=1}^{|\Pi|} p_{t-1}(j)(1 - \epsilon l(a_j, z_t))}$$

where $\epsilon \in (0, 1/2]$ and $l(a, z) \in [-1, 1]$. For gains we simply take the losses to be negative gain, and then the updates become

$$p_t(i) = \frac{p_{t-1}(i)(1 + \epsilon g(a_i, z_t))}{\sum_{j=1}^{|\Pi|} p_{t-1}(j)(1 + \epsilon g(a_j, z_t))}$$

The regret bound for MWU is [6]:

Theorem 6. *Assume all losses $l(a_i, z_t) \in [-1, 1]$ and $\epsilon \in (0, 1/2]$, then the Multiplicative Weight Updates algorithm guarantees that after T rounds, for any decision i , we have:*

$$\sum_{t=1}^T \sum_{i^*} l(a_{i^*}, z_t) p_t(i^*) \leq \sum_{t=1}^T l(a_i, z_t) + \epsilon \sum_{t=1}^T |l(a_i, z_t)| + \frac{\log n}{\epsilon}$$

And for MWU with gains the same bound holds, with the LHS lower-bounded by the RHS and losses replaced by gains [6]. From the above bound it follows that MWU is no-regret.

Recent work [25] has shown that infinite population selection with recombination and no mutation (and fitness values that are very close together, a regime called weak selection) is equivalent to MWU when used to play a coordination game (for a version of MWU that uses gains instead of losses). If one takes the limit of the same population genetics equations as used in [25] with no recombination (asexual reproduction) then one obtains

$$p_i^{t+1} = \frac{w_i}{\bar{w}_t} p_i^t$$

where w_i is the fitness and \bar{w}_t is the average fitness according to distribution p_i^t . When $w_i \in [1 - s, 1 + s]$ with $s \rightarrow 0$ (the weak-selection regime), and $\Delta_i = \frac{w_i - 1}{s}$ gives us:

$$p_i^{t+1} = \frac{1}{\bar{w}_t} p_i^t (1 + s \Delta_i)$$

which is just MWU with gains. We conclude that

Lemma 7. *Infinite population asexual reproduction with weak selection is MWU with gains.*

Now we will describe recent work giving an information-theoretic model of evolution,

and show how certain equations from population genetics are related to that family of models.

7.5 Process Information and Evolution

Evolution can be viewed as a process of communication between organism and environment [130, 120, 121, 119, 77]. Rivoire and Leibler sought to model evolution in a changing environment, and derive some relations between that setting and information theory including both the quantities, such as mutual information, and the mathematical formalism behind the theory. In the Rivoire and Leibler model [120], σ_t is the phenotype of which there are n_σ , x_t is the state of the environment of which there are n_x , y_t is a noisy signal of x_t , and $f(\sigma_t, x_t)$ is the multiplication rate (the expected number of offspring, or “fitness function”) for phenotype σ_t in environment x_t . In one special case of the model, $f(\sigma, x_t) = f(x_t)$ when $x_t = \sigma_t$ and 0 otherwise, and it is said to be diagonal. This special case refers to a situation in which each organism is capable of having a non-zero fitness only when its phenotypic state is equivalent to the state of the environment. Since organismal states are empirically almost always different from environmental states, the case of diagonal f is then considered to be rare as opposed to common. Even under a more abstract interpretation that treats the states as indices into sets of states with different meanings, this interpretation implies that an environmental state is associated to a single viable phenotypic state, which is also violated empirically. As the environment in the Rivoire-Leibler model is time-varying, it uses a discrete-time and discrete-state Markov chain to model how x_t varies over time, and further assumes its ergodicity and stationarity. The inherited information is the phenotype σ_{t-1} . Note in what follows we also call the time-evolution of the quantities in the Rivoire-Leibler model the Rivoire-Leibler process.

The strategy for changing the phenotype, called in information theory a communication channel [30], is transition matrix $\pi(\sigma_t|\sigma_{t-1}, y_t)$, with $\pi(\sigma_t|\sigma_{t-1}, y_t) \geq 0$ for all $\sigma_t, \sigma_{t-1}, y_t$, and $\sum_{\sigma_t} \pi(\sigma_t|\sigma_{t-1}, y_t) = 1$. Such a transition matrix is a communication channel with input (σ_{t-1}, y_t) and output σ_t . It is worth noting that π is a dynamics for updating the frequency of the phenotypes σ . The environment state x_t gives rise to a noisy cue

x'_t according to the communication channel $q_{env}(x'_t|x_t)$ and the cue gives rise to the noisy signal y_t according to the channel $q_{in}(y_t|x'_t)$. Noiseless channels q_{in} give perfect information about the cue, and are represented by the identity transition matrix δ such that $\delta(y_t|x_t) = 1$ if $y_t = x_t$, and 0 otherwise.

For example, let us consider a population of bacteria. x_t would give the chemicals constituting the medium at time t , x'_t the subset of those chemicals for which the bacteria have a sensor, and y_t the chemicals that a bacterium actually detects at time t , which may vary from bacteria to bacteria due to imperfect sensors. The difference between x_t , the state affecting the multiplication rate $f(\sigma_t, x_t)$, and x'_t , the cue, could model the delay between sensing and reproduction.

The fitness is a long-term growth rate for the number of organisms with the same phenotype given a particular π . In order to analyze how the population changes over time, Rivoire and Leibler introduce a vector Z_t with $Z_t(\sigma)$ being the frequency of individuals of type σ in the population, and the norm $|Z_t| = \sum_{\sigma} Z(\sigma)$ being the total population size. Z_t is a random vector, as it depends on the sequence $\bar{x} = ((x_1, x'_1), \dots, (x_t, x'_t))$, and given \bar{x} it is also subject to randomness generated by transition matrices π and p_{in} . To represent the average conditionally to the environmental sequence \bar{x} they use $\langle Z_t(\sigma) \rangle$, and $E[\langle Z_t \rangle]$ for the average over environmental sequences. Let $N_t(\sigma) = \langle Z_t(\sigma) \rangle$ be the average taken for a given \bar{x} . Then the following recursion holds:

$$N_t(\sigma_t) = f(\sigma_t, x_t) \sum_{\sigma_{t-1}, y_t} \pi(\sigma_t|\sigma_{t-1}, y_t) p_{in}(y_t|x'_t) N_{t-1}(\sigma_{t-1})$$

The recursion can also be written using linear algebra as $N_t = A^{(t)} N_{t-1}$, with

$$A_{\sigma'\sigma}^{(t)} = f(\sigma_t, x_t) \sum_{y_t} \pi(\sigma_t|\sigma_{t-1}, y_t) p_{in}(y_t|x'_t)$$

where $A^{(t)}$ is shorthand for $A^{(x_t, x'_t)}$, and the current environment (x_t, x'_t) is fixed independently of the dynamics π of the population. This is to say that $A^{(t)}$ is, in effect, a stochastically-chosen matrix whose randomness is induced by the stochasticity of

the environment. We initially make the following assumptions, which we subsequently relax:

1. No information is inherited between generations, $\pi(\sigma_t|\sigma_{t-1}, y_t) = \pi(\sigma_t, y_t)$
2. Any information acquired from the environment is common to all members of the population, so $p_{in} = \delta$, and thus $y_t = x'_t$.
3. The multiplication rates have a diagonal form (see above).

Then if our objective is to consider only the long-term growth of the total population size, the fitness can be defined as:

$$\Lambda_{q;f}^{(p_{env}, p_{in})} = \lim_{t \rightarrow \infty} \frac{1}{t} \log |N_t|$$

We want to choose π to maximize the growth rate for N_t . If one chooses $\hat{\pi}$ to maximize $\Lambda_{q;f}^{(p_{env}, p_{in})}$, then $\hat{\pi}$ outcompetes all other time-varying π 's [120], in other words, the following holds with probability 1:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \frac{|N_t(\pi)|}{|N_t(\hat{\pi})|} \leq 0. \quad (7.2)$$

Rivoire and Leibler quote a result by Kelly which implies that with probability 1,

$$\Lambda_{q;f}^{(p_{env}, p_{in})} = \lim_{t \rightarrow \infty} \frac{1}{t} E[\log |N_t|]$$

where

$$\lim_{t \rightarrow \infty} \frac{1}{t} E[\log |N_t|] = \sum_{x, x'} p_{env}(x'|x) p_s(x) \log \left(\sum_{\sigma, y} f(\sigma', x) \pi(\sigma'|y) p_{in}(y|x') \right)$$

and $p_s(x)$ is the stationary probability of the state x for the Markov chain modeling the states of the environment. If the environmental state is drawn i.i.d from D instead of from a Markov chain, then instead of $p_s(x)$ we will have $p_D(x)$.

The intention of the model is to show that the value of the information conveyed by p_{env} and p_{in} is related to information-theoretic quantities. A way of modeling the

value of information is

$$\hat{\Lambda}_{q;f}^{(p_{env}, p_{in})} - \hat{\Lambda}_{q;f}^{(\nu, \nu)}$$

where $\hat{\Lambda}$ means the fitness for $\hat{\pi}$ and ν is an “informationless channel” $\nu(y|x) = *$ where $*$ signifies no output. It is shown that $\hat{\Lambda}_{q;f}^{(p_{env}, p_{in})} - \hat{\Lambda}_{q;f}^{(\nu, \nu)} = I(X_t, Y_t)$ where $I(X_t, Y_t)$ is the mutual information between the state of the environment (X_t is a random variable taking values x_t) and the signal (Y_t is a random variable taking values y_t). They furthermore show that the reduction of uncertainty $\hat{\Lambda}_{q;f}^{(\delta, \delta)} - \hat{\Lambda}_{q;f}^{(p_{env}, p_{in})}$ (where δ is a noiseless channel) is equal to $H(X_t|Y_t)$. They furthermore show that there are generalizations of these quantities that upper-bound the value of information and reduction in uncertainty when Assumptions (1)-(3) do not hold.

It has been shown however that one of the special cases of Equation 7.2 (when the state x_t is drawn i.i.d from some distribution) is merely the regret when using the log-loss function [23]. The optimal estimate $\hat{\pi}$ is solved for by Kelly and Cover explicitly [71, 30], and quoted by Rivoire and Leibler. But the optimality measure given by the equation is regret, and indeed, if one attains 0 regret as $t \rightarrow \infty$ with probability 1, one has a solution that is of the same quality as $\hat{\pi}$ since it is only optimal in the limit as $t \rightarrow \infty$. So a no-regret estimate $\hat{\pi}$ is what Rivoire and Leibler give an approach for finding. The relation between no-regret and optimal prediction has been shown in [23], interpreting classic results by Cover on log-optimal portfolios. Thus we can use the connection to prove the following:

Theorem 8. *Assuming the actions $a \in \Pi$ are elements of the n_x simplex for finite set Π , the Rivoire-Leibler model gives a no-regret protocol π with probability 1, when Nature selects z_t according to an i.i.d process with $l(a, z) = -\log(a^T z)$ being the loss for each action and z having exactly one non-zero component which is equal to $f(x_t)$ for a fixed function known to the learner.*

One can alternatively give a no-regret sequence of π that vary over time, and such a sequence can be generated by the multiplicative weight updates, regardless of whether the environment is Markov which we will show. MWU applied to this setting will have a finite set Π of a 's, with $\hat{\pi} \in \Pi$, and is given by the following protocol:

$$p_t(i) = \frac{p_{t-1}(i)(1 + \epsilon \log(a_i^T z_{t-1}))}{\sum_{j=1}^{|\Pi|} p_{t-1}(j)(1 + \epsilon \log(a_j^T z_{t-1}))}$$

for all $i \in |\Pi|$, where $p_t(i)$ is i^{th} index of the strategy and p_t is the probability vector for choosing the different actions in Π . This update can either be viewed as the gains version of MWU or the loss version with the log-loss $-\log(a^T z)$. The connection between the Rivoire-Leibler model and online learning is not only interesting as a connection between regret and long-term measures of fitness, but also gives a connection between the Rivoire-Leibler model of evolution and population genetics. As a simple corollary of Theorem 8 and Lemma 7, it follows that infinite population asexual haploid selection with no mutation has comparable growth rate (e.g. both have comparable long-term fitness) to the population with the initial probability for the experts set to $p_0(x) = p_s(x)$.

Theorem 9. *There exists a set of fitnesses of the form $w_i = -\log(a_i^T z)/M$ such that Infinite population asexual selection with no mutation gives a π with $\Lambda_{q,f}^{(p_{env}, p_{in})} - \hat{\Lambda}_{q,f}^{(p_{env}, p_{in})} \leq 0$ with probability 1 when Assumptions (1)-(3) hold and the state of the environment is chosen i.i.d.*

Proof. See Section 7.6 □

And so Rivoire and Leibler’s model, stated in the abstract, actually describes the behavior of a classic model in population genetics. We can also take these dynamics and run them on non-diagonal multiplication factors (drop Assumption (3)). The closed form for the non-diagonal multiplication factors is not given by Rivoire-Leibler and seems difficult to derive. But we don’t have to solve for the optimal estimate! The good news is that we can have an estimate which is asymptotically accurate but based on something other than the optimal estimate. Such an estimate will be the set of probabilities $p_t(i)$ over time given by MWU. We will detail how one could use MWU to calculate an asymptotically-optimal π . But first we will discuss why exactly non-diagonal multiplication rates are so important.

According to Rivoire and Leibler, “When $f(\sigma_t, x_t)$ isn’t diagonal, the environmental states have no longer an exclusive meaning, in the sense that the same environment

can be beneficial to different types, and different environments to the same type.” Of course this situation happens quite often. Many examples exist, but one is the fact that seven distinct *Anolis* lizard species share common food needs (mainly insects).

More generally, this phenomenon gives rise to competition over food resources, with implications for multiplication rates. Non-diagonal fitness functions also arise in a frequency-dependent fitness landscape, in which for instance the fitness of phenotype σ is dependent on the current frequency $p_t(\sigma')$ of phenotype σ' in the population [125]. For instance, if the fitness of σ is proportionate to the frequency $p_t(\sigma')$ at time t of σ' in the population and vice-versa (called positive frequency dependence), then $f(\sigma, x_t) = kp_t(\sigma')$ and $f(\sigma', x_t) = k_2p_t(\sigma)$ and thus generically over frequencies of both σ and σ' the fitnesses are both positive at the same time, and thus $f(\sigma, x_t)$ has positive off-diagonal terms for positive frequency-dependence when it has positive diagonal terms.

In the following result, we show that for non-diagonal multiplication rate one can use MWU to make an online learning protocol which is no-regret as compared to the optimal growth rate, thus by Theorem 8 giving a protocol for this case that improves the Rivoire-Leibler result. Moreover, our result not only holds for the Markov chain model of the environment, but also could hold for adversarially-chosen environments (if one suitably generalizes the definition of long-term fitness). We do not provide this generalized long-term fitness for the sake of space constraints, but it is not too difficult to derive. The key insight for this result is that z for non-diagonal multiplication rate has more than one positive entry:

Theorem 10. *There exists a set of fitnesses of the form $w_i = -\log(a_i^T z)/M$ such that Infinite population asexual selection with no mutation gives a π with $\Lambda_{q;f}^{(p_{env}, p_{in})} - \hat{\Lambda}_{q;f}^{(p_{env}, p_{in})} \leq 0$ when Assumptions (1)-(2) hold for the Rivoire-Leibler process.*

Proof. See Section 7.6 □

The environmental states having the same “meaning” due to equality of multiplication rate seems to be rooted in an intuition that if one can’t distinguish the two with respect to information, then they are the same. This at its core is an intuition about

information. According to papers on functional information, a proper measure of information for biology should take into account that different RNA sequences can have the same function. That is, there is tremendous functional redundancy (for instance, in RNA sequences and things like the synthesis of catalysts and aptamers) [133]. To model this, one would have to use relational information, and in particular, synonymy to the same concept/function. We quoted Rivoire and Leibler before on how they would approach this. Let's flesh their model out a bit:

“The source of meaning, encapsulated in the values of the multiplication rates of the individuals, needs to be taken explicitly into account in the measure of information”

For this to make sense, we have to couple the multiplication rate of individuals to the meaning somehow (besides just defining it to be the values of the multiplication rates themselves).

Let us assume that rather than $f(\sigma, x)$ giving us only a multiplication rate, it gives us a measure of compatibility, by which we say σ is compatible with x to some degree (as was suggested by Rivoire and Leibler). Recall that the concept-word feedback function from process information $r_c(\sigma, \sigma', x)$ plays a similar role, giving a measure of compatibility between σ and x , but also quantifying to what degree σ' is just as compatible as σ to x . Say $r_c(\sigma, \sigma', x)$, when nonzero, gives a measure of how compatible both are to x , and if $r_c(\sigma, \sigma', x) = 0$ then both are not equally compatible. Then when $r_c(\sigma, \sigma', x) > 0$, $f(\sigma, x) = f(\sigma', x) = r_c(\sigma, \sigma', x)$ and otherwise $f(\sigma, x) \neq f(\sigma', x)$. Then by the definition of r_c , equality of multiplication rates for different phenotypes is entirely determined by whether σ and σ' considered as sets in some universe C have associated functions that have $x \in C$ as a codomain. Now recall that the pullbacks are just the sets which make each of these equal. So the pullbacks are sets of all equal meaning phenotypes and states. Then the pullbacks are themselves sets and we can apply a function to these sets to get the equal meaning phenotypes and the equal meaning states. What this shows is that the equal multiplication rate condition corresponds to finding a form. We will use this insight algorithmically.

We will show that there is a way to use maximization of growth rate to test whether a pair of phenotypes σ, σ' satisfy $f(\sigma, x) = f(\sigma', x)$. Then we show that this can be

used to predict the pullbacks of all n phenotypes with only $\text{poly}(n)$ different evolving populations. Thus the evolving population of phenotypes in a Rivoire-Leibler model will be able to decipher which phenotypes are compatible with which states of the environment, and whether they are equally so. The evolving population according to Rivoire and Leibler's model thereby possesses process information relevant to the environment.

Theorem 11. *Bob can be informed, in the sense of process information, using $O((|A| + |B|)^2|\psi|)$ Rivoire-Leibler process optima for a Rivoire-Leibler process with i.i.d states of the environment, drawn uniformly at random. Furthermore the algorithm used is conservative.*

Proof. See Section 7.6. In outline, we first show how one can use a Rivoire-Leibler growth rate maximum $\hat{\pi}$ to decipher whether a pair of phenotypes have the same multiplication rate for a fixed environment. Then we show how to use this to recover all the pullbacks of the set. Finally, we show, in particular, that this can be accomplished using $O((|A| + |B|)^2|\psi|)$ Rivoire-Leibler process optima. \square

Based on the connection between MWU and the Rivoire-Leibler model optima, and using a similar argument as the previous results (Theorem 10) showing that the population genetics dynamics attain the Rivoire-Leibler growth optima for some fitness function, we obtain:

Theorem 12. *Bob can be informed, in the sense of process information, using $O(|A| + |B|)^2|\psi|)$ runs of no-mutation infinite population asexual selection dynamics for a Rivoire-Leibler process with i.i.d states of the environment, drawn uniformly at random.*

Then Bob can be $m(n)$ -informed, where $m(n) = O(4n^3)$, since $(2|U|)^2|U| = O(|A| + |B|)^2|\psi|)$. The algorithm used was conservative. Thus we can conclude by using Theorem 5 that:

Theorem 13. *Infinite population asexual selection with no mutation (on the appropriate inputs) gives a $(U, 1, 1)$ -generic universal user for G_C for error bound $m(n(U)) + 1$ where $m(n) = O(4n^3)$.*

Finally, Rivoire and Leibler’s model has been applied as a justification for making information-theoretic models of biological systems, as then the fitness function is in some sense proportionate to information theoretic quantities (see, for example, [79]). But the foundations for this identification are unclear, since in order to apply the analysis we have to have some idea about what the information processed by the system is about [134]. Thus one could use the generalization of the Rivoire-Leibler model we have proposed here to give a better foundation for the use of information theory in modeling biological systems, by using a version of communication theory that takes semantics into account. We leave a more detailed pursuit of this goal for future work.

Acknowledgements We thank Eric Allender and Lucy Chastain for comments on the manuscript. We thank Raymond Puzio, Jay Sulzberger, Ximo Pechuan, and Aviv Bergman for helpful discussions. Support was provided by NSF grant CCF-1555409 to Erick Chastain, and NIH MSTP training grant T32-GM007288 to Cameron Smith.

7.6 Proofs of Evolutionary theory and MWU Results

In this section we give the proof of major results relating Evolutionary theory, the Rivoire-Leibler model, and MWU.

Proof of Theorem 9. Without loss of generality, we can re-scale $f(x)$, the multiplication rate or fitness function, to be in the interval $[-1, 1]$ by dividing by $\max_x |f(x)|$, and since Π is finite, we can rescale the log-loss, dividing it by the maximum of $-\log(a^T z)$ over all $a \in \Pi$ to be in the interval $[-1, 1]$. Then we use Lemma 7 on the resulting fitness function (which is in the weak-selection regime) and apply Theorem 6, therefore obtaining that the resulting dynamics are no-regret. When we take the time-average of all the losses, asymptotically the resulting quantity is exactly $E[\log(a^T z)]$ w.p. 1, by the weak law of large numbers, so by showing that the dynamics are no-regret, we have shown that $\Lambda_{q,f}^{(p_{env}, p_{in})} - \hat{\Lambda}_{q,f}^{(p_{env}, p_{in})} \leq 0$ asymptotically w.p. 1 for the dynamics.

□

Proof of Theorem 10. First we will show that we can construct a fitness function for the

population genetics dynamics such that MWU on the appropriate losses is simulated. Then we will show that when we use actions $a \in \Pi$ such that Π is a finite set of elements of the n_x simplex, MWU gives an optimal growth rate learning protocol π with probability 1, when Nature selects z_t according to an i.i.d process for $\lambda = (x, x')$ for $p(\lambda) = p_{env}(x'|x)p(x)$ with $l(a, z) = -\log(a^T z)$ being the loss for each action and z having components which are equal to $f(a, x_t)p_{in}(y|x')$ for a fixed function known to the learner.

For each y_t we can do a separate evolution process (which corresponds to MWU for each y_t), then we will have an optimal growth population conditioned on each y_t , and thus for all y_t .

So conditioned on a particular y_t : Without loss of generality, we can re-scale $f(\sigma, x)$ to be in the interval $[-1, 1]$, and since Π is finite, we can rescale the log-loss, dividing it by the maximum of $-\log(a^T z)$ over all $a \in \Pi$ to be in the interval $[-1, 1]$. Then we use Lemma 7 on the resulting fitness function (which is in the weak-selection regime) and apply Theorem 6, therefore obtaining that the resulting dynamics are no-regret. When we take the time-average of all the losses, asymptotically the resulting quantity is exactly $E_{p(\lambda)}[\log(a^T z)]$ w.p. 1, by the weak law of large numbers, so by showing that the dynamics are no-regret, we have shown that $\Lambda_{q;f}^{(p_{env}, p_{in})} - \hat{\Lambda}_{q;f}^{(p_{env}, p_{in})} \leq 0$ asymptotically w.p. 1 for the dynamics.

□

Proof of Theorem 11. The sets of Alice are some subset P_A of phenotypes, and the sets of Bob are some subset P_B of phenotypes, and the concept sets are the states S_C of the environment. Let the set of all phenotypes P be $P = P_A \cup P_B$. Let the feedback function r_c give the multiplication rate function $f(i, x)$ as follows: the words are phenotypes (as we saw above) and the states of the environment are concepts, so $f(w, c) = f(w', c) = r_c(w, w', c)$ if $r_c(w, w', c) > 0$ and $f(w, c)$ is set as follows: Let $O_{ij} = f(j, i)$ be the matrix of multiplication rates, and D , with $D_{ij} = \delta_{ij}(\sum_k (O^{-1})_{ki})^{-1}$. Set the entries of O so that D is a positive matrix. We will see an example of such a case below.

We will first show how one can use a Rivoire-Leibler growth rate maximum $\hat{\pi}$ to

decipher whether a pair of phenotypes have the same multiplication rate for a fixed environment. Then we will show how to use this to recover all the pullbacks of the set.

Let $O_{ij} = f(j, i)$ be the matrix of multiplication rates, D , with $D_{ij} = \delta_{ij}(\sum_k (O^{-1})_{ki})^{-1}$ be a positive diagonal matrix, and p be a vector containing the probability $p(x)$ of the state of the environment x for all x . Then the optimal growth rate estimate $\hat{\pi}$ assuming that $y_t = x_t$ is given by $\hat{\pi} = O^{-1}Dp$ [116]. We will use the above formulas to create a gadget for the reduction.

To figure out whether for a state x a pair of phenotypes i and i' are equal in multiplication rate, one can make the matrix O' of multiplication rates for a Rivoire-Leibler process which has entries

$$O' = \begin{bmatrix} f(i, x) & b \\ f(i', x) & f(i, x)^2/b \end{bmatrix}$$

where b will be chosen in accordance with D being positive, which gives us $f(i, x) > b > 0$, and all positive entries.

Without loss of generality, let's assume $f(i, x) \neq 0$ (one can re-scale all of the multiplication rates additively and re-do the argument accordingly). For convenience, let $a = f(i, x)$, $c = f(i', x)$ and $d = a^2/b$. Then $D_{11} = 1/d - c$ and $D_{22} = 1/a - b$.

Since we assumed in the reduction that the entries of O were set such that D is positive, $d > c > 0$, $a > b > 0$. This ensures that $c \neq d$ even when $a \neq c$, which preserves $ac = bd$ as the key criterion for the equality test to succeed, and that at least as long as a is nonzero, $a - b$ is also nonzero, and furthermore when $a = c$, $c - d$ is also nonzero. All of these properties are important for the positivity of $O'^{-1}Dp$ and for the validity of the test we propose for testing equivalence of growth rate.

Now the matrix O' is a 2 by 2 matrix that when $a = c$ has (by the above inequalities given by the positivity of D):

$$\det(O') = ad - bc > 0$$

When $f(i, x) = f(i', x)$, O'^{-1} is given by:

$$O'^{-1} = \frac{1}{a(d-b)} \begin{pmatrix} d & -b \\ -f(i', x) & f(i, x) \end{pmatrix}$$

Note that, $Dp = \langle D_{11}p_1, D_{22}p_2 \rangle$

Take the Rivoire-Leibler process to have multiplication rate matrix O' , and to have x_t be drawn uniformly at random, then $p_i = p_j$ for all i, j , and we obtain:

$$O'^{-1}Dp = [dD_{11}p_1 - bD_{22}p_2, aD_{22}p_2 - cD_{11}p_1]/\det(O')^2$$

$$O'^{-1}Dp = [p_1(dD_{11} - bD_{22}), p_1(aD_{22} - cD_{11})]/a^2(d-b)^2$$

$$O'^{-1}Dp = [p_1(d/(d-c) - b/(a-b)), p_1(a/(a-b) - c/(d-c))]/a^2(d-b)^2$$

$$\hat{\pi}_{(x,i,i')}(1) = p_1(d/(d-c) - b/(a-b))/a^2(d-b)^2$$

$$\hat{\pi}_{(x,i,i')}(2) = p_1(a/(a-b) - c/(d-c))/a^2(d-b)^2$$

And note that since $a(d-b) > 0$ (and thus $a^2(d-b)^2 > 0$), these probability vectors are positive, as required. It turns out that if

$$\frac{1}{\det(O')^2}[d/(d-c) - b/(a-b)] = \frac{1}{\det(O')^2}[a/(a-b) - c/(d-c)]$$

if and only if

$$ac = bd$$

But when $a = c$, then $ac = a^2$ and $bd = b(a^2/b) = a^2$ (as $b \neq 0$), and therefore $ac = a^2 = bd$. Thus when $a = c$, $\hat{\pi}_{(x,i,i')}(1) = \hat{\pi}_{(x,i,i')}(2)$.

So if the multiplication rate is the same between phenotypes in environmental state x , $\hat{\pi}(2) = \hat{\pi}(1)$.

Let the equality test “succeed” if $\hat{\pi}_{(x,i,i')}(1) = \hat{\pi}_{(x,i,i')}(2)$ and both $\hat{\pi}_{(x,i,i')}(1)$ and $\hat{\pi}_{(x,i,i')}(2)$ exist.

Let’s check whether the equality test implies that $f(i, x) = f(i', x)$. If the asymptotic probabilities exist (that is, O' is invertible), then:

$$\hat{\pi}_{(x,i,i')}(1) = p_1(d/(d-c) - b/(a-b))/\det(O')^2$$

$$\hat{\pi}_{(x,i,i')}(2) = p_1(a/(a-b) - c/(d-c))/\det(O')^2$$

Note that these probability vectors too are positive, since as we showed above, $\det(O') = ad - bc > 0$ and thus $\det(O')^2 > 0$.

Now $\hat{\pi}_{(x,i,i')}(1) = \hat{\pi}_{(x,i,i')}(2)$ implies that $d/(d-c) - b/(a-b) = a/(a-b) - c/(d-c)$ and thus $ac = bd$ for $a \neq 0$.

We will now prove by contradiction that the equality test, if it succeeds, would never imply that $a \neq c$:

Assume that the equality test succeeds and $a \neq c$. The equality test succeeds only if $ac = bd$. But $bd = a^2$, by the definitions of b and d , since $b \neq 0$. Therefore, $ac = a^2$, and thus $c = a$. But we assumed that $a \neq c$. Thus the claim holds.

Therefore if the equality test succeeds, the only way it could succeed is if $a = c$ and thus $f(i, x) = f(i', x)$.

Now when we run the Rivoire-Leibler process with the above parameters, if we run the equality test and it succeeds, at any entry, this is a test for sharing the same form for i and i' given some state of the environment x . Say that this test is denoted as $\Psi(\hat{\pi}_{(x,i,i')})$ and outputs a set of 3-tuples of the form (i, i', x) such that $f(i, x) = f(i', x)$. We can run a different process for each state x in the set S of environmental states and obtain the set of estimates for the pair of phenotypes i, i' : $F(i, i') = \{\cup_{x \in S} \Psi(\hat{\pi}_{(x,i,i')})\}$. Finally, we can run these Rivoire-Leibler processes for other possible pairs of phenotypes besides i and i' in the set of possible phenotypes P . Then we take $F = \cup_{i, i' \in P} F(i, i')$. Then F is a set of all forms associated with the phenotypes in P and the states in S . Clearly the algorithm to calculate F only eliminates tuples when $r_c(w, w', c) = 0$ and is

thus conservative. To calculate F , we must run $\binom{|P|}{2}|S|$ Rivoire-Leibler processes. This simplifies to $O(n_\sigma^2 n_x)$ Rivoire-Leibler processes, which gives the result, since we can use F to give the forms for all phenotypes and $n_\sigma = |P| = |P_A \cup P_B| \leq |P_A| + |P_B|$.

□

7.7 Proofs of Process information and USC Results

What follows is a more detailed brief summary of USC followed by a proof of Theorem 5. For more information on the particular version of USC we use, see [69]. The table that follows summarizes all of the terms used in the model:

notation	meaning
U	user strategy set
S	server strategy set
E	non-deterministic environmental model
k	session length
E_1	start-session states
E_k	end-session states
$R' : E_k \rightarrow \mathbf{2}$	temporal decision function evaluating end-session states
R	referee determines verdict according to R'
G	goal = (E, R)
Ω	state space $X \equiv (u, s, e) \in \Omega$ for some $u \in U, s \in S, e \in E$
X_1, \dots, X_t, \dots	stochastic process over Ω

Recall that USC has a user and a server accomplish some goal together while interacting with their environment. User, server, and environment each have some internal state, and they are each joined by a (two-way) communications channel that also has a fixed state on each round. User, server, and environment have a strategy that specifies a distribution over new internal states and outgoing messages for the following round, given the entity's current state and incoming messages. Given strategies for each of the entities, the system as a whole is modeled by a discrete-time Markov process with a state space Ω . We call an execution the infinite sequence of random variables X_1, \dots, X_t, \dots

where X_t is the state of system in round t . An execution produced by the interaction between a user strategy u , a server strategy s , and an environment strategy e will be denoted by (e, u, s) . An execution started from state σ_1 is an execution condition on $X_1 = \sigma_1$. We denote the internal states of the user, server, and environment by $\Omega^{(u)}$, $\Omega^{(s)}$, and $\Omega^{(e)}$ respectively, and for $i, j \in \{u, e, s\}$ the state of the communication channel from i to j is a member of $\Omega^{(i,j)}$. Given a state of the system σ , we will let the respective superscripts denote the projection of σ onto the respective components (so $\sigma^{(s,e)}$ is the server's outgoing message to the environment in σ).

A goal $G = (E, R)$ where E is a non-deterministic environmental model and R is a referee that classifies whether the goal is achieved. A user strategy $u \in U$ is S -universal with respect to a goal G if for every server strategy $s \in S$ (u, s) robustly achieves the goal (achieves it for all $e \in E$ and from all initial conditions). We will discuss how to assure the robust achievement of a goal when we discuss sensing functions.

A goal $G = (E, R)$ is said to be a k -round multi-session goal if the following hold:

1. The system's states Ω are partitioned into k sets $\Omega_1^{(e)}, \dots, \Omega_k^{(e)}$. $\Omega_1^{(e)}$ are the start-session states, and $\Omega_k^{(e)}$ are the end-session states. Each of the states is a pair consisting of an index and a contents.
2. When in an end-session state, the environment (non-deterministically) moves to a start-session state with index incremented (and the target state is independent of the end-session state).
3. When the environment is in some state $(j, \sigma) \in \Omega_i^{(e)}$ for $i \neq k$, $E(j, \sigma)^{(e)}$ is a distribution over $\Omega_{i+1}^{(e)}$ with every element in its support having index j . The distribution is independent of the index and the environment's strategy.
4. There is a temporal decision function R' giving boolean verdicts for end-session states, and R is satisfied with an infinite execution iff R' evaluates to zero at most finitely many times, and the number of times this happens is called the number of errors.

Let $G = (E, R)$ be a fixed-length multi-session goal with temporal decision function

R' and size parameter function $sz : \Omega \rightarrow \mathbb{N}$, let $s \in S$ be a server strategy, and U be a class of user strategies. Sensing functions in USC are boolean functions taking all events in a series of interactions between the server and user and giving goal-specific feedback. In USC, one has v -viability for G for a sensing function V if there is a user strategy $u \in U$ such that for all $e \in E$, $\sigma_1 \in \Omega$, V after $v(sz(\sigma_1))$ rounds evaluates to 1 in every subsequent round in the execution of the protocol (where σ_1 is an initial state used by the execution of the protocol) with probability 1. V is s' -safe for G if for all $e \in E$, user strategies $u \in U$, starting states σ_1 , whenever $R'(\sigma_1) = 0$, then at most $s'(sz(\sigma_1))$ errors will occur, or for some $t \leq s'(sz(\sigma_1))$, V evaluates to 0 in some state X_t of the execution (e, u, s) started from σ_1 with probability 1.

Definition 17. For a class of goals in infinite executions G , a class of user strategies U , and bound $B : U \times \mathbb{N} \rightarrow \mathbb{N}$ taking a user strategy and a target error rate and giving the number of rounds necessary to attain that, we say that $u \in U$ is a B -error (U, s', v) -generic universal user for G if for all $g \in G$, any server s , and any sensing function V that is s' -safe and v -viable with s with respect to U for G when $u \in U$ is provided the verdicts of V as auxiliary input, (u, s) robustly achieves G with $\min_{U_S \in U: U_S \text{ } v\text{-viable with } s} B(U_S, \cdot)$ errors.

Proof of Theorem 5. U is the set of functions mapping from a word to its signifier. Suppose we are given white box/full access (not just black-box input/output access, but access to all internal variables) to the operations of a $m(n)$ -informing process information algorithm A for U . Again, and more specifically, we assume that the simulation of the USC protocol involves using and setting variables / functions internal to the algorithm itself during its execution. We show that A is a generic universal user. Suppose that we are given a goal $G \in G_C$, a server $s \in S$, and a sensing function V that is 1-safe with s for G and 1-viable with s with respect to U for G .

By the 1-viability of V , there is a $U_S \in U$ s.t. if the user sends the same messages as U_S , after one round V will give a positive indication on every round.

Each round of the execution corresponds to a triplet (w, w', x) for the process information algorithm A . Recall that pairs of words (w, w') are given to A each round and

then the input to r_c includes w, w' and an associated state x .

Suppose we provide the incoming messages to A as the words w, w' for each triplet (providing to A a pair of words (w, w') rather than some other pair), and take the outgoing messages as the signifier x of A 's associated triplet (as determined by A 's chosen x to be used in any call of $r_c(w, w', x)$ itself associated with the desired pair of words (w, w')), and provide the verdict of V on the following round as the reinforcement function $r_c(w, w', x)$ for that round.

If A 's associated triplet (w, w', x) contains the same x as is sent in the outgoing message as U_S , the sensing function will give to the triplet of A a positive indication, which we take as positive feedback for the triplet associated with A . If V produces a negative indication, then the triplet associated with A must not have contained the same x as U_S would have sent on the incoming messages in that round. V may also produce positive indications when the outgoing message from A 's associated triplet differs from what U_S would have sent, but since A is conservative, the state of A does not change. Now, since A is a $m(n)$ -informing algorithm for U , it only receives negative reinforcement $m(n)$ times in any execution.

As G_C is a 1-round multi-session goal, R' becomes 0 or 1 on each round; when 0, the 1-safety of V implies either that is the only error that will occur, or that V evaluates to 0 in the current round. V thus can only evaluate to 1 if an error occurs once, so the overall strategy makes at most $m(n) + 1$ errors. \square

7.8 Category Theory

A category C is given by a collection $Obj(C)$ of objects and a collection $Morph(C)$ of morphisms which have the following structure (see [89, 131] for a general introduction to category theory):

Each morphism has a domain and a codomain which are objects, writing $f : X \rightarrow Y$ or $X \xrightarrow{f} Y$ if X is the domain of the morphism f , and Y is its codomain. Also often used notationally is $X = dom(f)$ and $Y = cod(f)$. Given two morphisms f and g

such that $\text{cod}(f) = \text{dom}(g)$, the composition of f and g , written gf , is defined and has domain $\text{dom}(f)$ and codomain $\text{cod}(g)$:

$$(X \xrightarrow{f} Y \xrightarrow{g} Z) \mapsto (X \xrightarrow{gf} Z)$$

Composition is associative, so given $f : X \rightarrow Y$, $g : Y \rightarrow Z$, and $h : Z \rightarrow W$, $h(gf) = (hg)f$. For every object X there is an identity morphism $\text{id}_X : X \rightarrow X$ satisfying $\text{id}_X g = g$ for every $g : Y \rightarrow X$ and $f(\text{id}_X) = f$ for every $f : X \rightarrow Y$. A morphism $f : X \rightarrow Y$ in C is an isomorphism if there is another morphism $g : Y \rightarrow X$ in C such that $gf = \text{id}_X$ and $fg = \text{id}_Y$. Two objects X and Y are isomorphic if there exists an isomorphism f with $\text{dom}(f) = X$ and $\text{cod}(f) = Y$.

Example: **Set** is the Category in which $\text{Obj}(\text{Set})$ contains sets, and $\text{Morph}(\text{Set})$ contains functions.

Given morphisms $f : X \rightarrow Z$ and $g : Y \rightarrow Z$, the pullback of f and g consists of an object P and two morphisms $p_1 : P \rightarrow X$, $p_2 : P \rightarrow Y$ for which the diagram

$$\begin{array}{ccc} P & \xrightarrow{p_2} & Y \\ p_1 \downarrow & & \downarrow g \\ X & \xrightarrow{f} & Z \end{array}$$

commutes. In other words, $fp_1 = gp_2$. The pullback if it exists is unique up to isomorphism. For the category **Set**, the pullback is a set $(X \times_Z Y) = \{(x, y) \in X \times Y \mid f(x) = g(y)\}$ together with the projection maps $p_1 : (X \times_Z Y) \rightarrow X$ and $p_2 : (X \times_Z Y) \rightarrow Y$. We say that a category C has pullbacks if for all pairs of morphisms f and g in $\text{Morph}(C)$ there always exists a pullback.

7.9 Category-Theoretic Process Information

The background concepts from category theory we apply here are reviewed in Section 7.8. The basic generalization of the Process Information model is replaces functions with morphisms and sets with objects. The universe C is replaced by an ambient category C . Rather than using the projection functions and pullbacks defined for sets,

we use the more general notions from Category Theory given in Section 7.8.

We assume that both Alice and Bob's words are objects in the same category C which has pullbacks. Alice's words are $Obj(A) \subseteq Obj(C)$, and Bob's words are $Obj(B) \subseteq Obj(C)$. Alice and Bob are assumed to each have a collection of words associated to the same collection of concepts. The elements of this common collection of concepts are also taken to be objects in C as well, which are the objects $Obj(\psi) \subseteq Obj(C)$. Let $Q \subset Obj(A) \times Obj(B)$ be a set of (w, w') pairs that refer to or signify the same concept. To model Diagram 7.1 (for Alice and Bob's words and concepts), we assumed that the same concept mapped to multiple words $w \in Obj(A)$ and $w' \in Obj(B)$. We also model the other content in Diagram 7.1 by saying for all pairs of words $w \in Obj(A)$ and $w' \in Obj(B)$ and $(w, w') \in Q$ there exists some concept $c \in Obj(\psi)$ together with morphisms $f_A : w \rightarrow c$ and $f_B : w' \rightarrow c$. As there is at most one form for every pair of words (though not every word pair is associated with a form), one way of modeling this is to assume that the form $F(w, w')$ of w, w' is the pullback object of f_A and f_B , and since C has pullbacks over $Obj(C)$, $F(w, w')$ always exists for all objects in $(Obj(A) \cap Q) \cup (Obj(B) \cap Q)$. Moreover it is unique up to isomorphism, by the uniqueness of pullbacks up to isomorphism.

Chapter 8

Conclusion

We have shown in this thesis a variety of applications of the MWU method and other tools from theoretical computer science to deriving new models in evolutionary biology. Naturally, there are many issues and future directions that have been unearthed by this investigation. First we will discuss particular domain areas, then we will discuss more general issues.

8.1 Population Genetics and MWU

Les Valiant [137] was the first to point out a connection between Evolution and Learning: The class of traits which are achievable by a species through random mutations constitute a specialized kind of learnability (a subcase of statistical learning, actually). Here we point out a different connection between these two fundamental computational categories: On the Wright manifold (that is, when linkage disequilibrium can be disregarded), Evolution *is* in fact learning through multiplicative updates — a well-known learning algorithm of very broad appeal. The same reasoning establishes Evolution as a coordination game between genes. Genes are the players, alleles are the strategies, and the mixed strategies are the allele statistics in the population. Notice how the organism is sidelined in this viewpoint: Evolution is an interaction between the genes (acting in a seemingly strategic manner) and the population (which stores the system's state). This is consistent with the gene-centric view of evolution, by which much can be explained by just focusing on evolution of and interaction between genes rather than organisms or species [132].

There is an important difference between Valiant's work and our point of view:

Valiant's evolvability is concerned with understanding random mutations. We are concerned in the ways in which sex and recombination nudge allele frequencies in the direction of high mixability (expected fitness); mutations are only implicit in our model, they are the (much slower) background process that creates the diversity exploited by recombination. Evolution and coordination games are not to be thought of as identical, however. A key difference between Evolution and coordination games is this: In a coordination game, a player may switch to a strategy which is currently not in the player's support (is being played with probability zero), once it becomes favorable. In Evolution, once an allele becomes extinct it never comes back (except through a new mutation, at a time scale far bigger than our current concerns). That is why the equilibrium of the process may not be a Nash equilibrium of the original game (but it is a mixed Nash equilibrium of the subgame defined by the support).

Starting with Darwin, researchers have often expressed an instinctive sense of disbelief that all Life we see around us could have come about through the rather rudimentary processes envisioned by Evolution. Connecting Evolution with multiplicative updates may help a little in this cognitive/cultural difficulty; after all, the multiplicative updates algorithm has surprised us time and again with its seemingly miraculous performance and applicability.

This work leaves open a variety of questions. Besides weak selection, for which other classes of fitness landscapes is the present analysis applicable? It is known that *product landscapes* (the fitness of a genotype is the product of fitness values, one for each allele present in the genotype) have the property of *staying* on the Wright manifold, once they are started there, but are not in general attracted to it [20]. It would be very interesting to come up with a combinatorial characterization of the fitness landscapes that converge to the Wright manifold (like the weak selection landscapes) and of those which at least stay there (like the multiplicative landscapes).

Secondly, in the context of the diversity proof in Chapter 4, consider the following computational problem: "flip the rows and columns of a given nonsingular square matrix such that the inverse has nonnegative row and column sums." The existence proof through a potential function places this problem in the class PLS [65]. Is it

PLS-complete?

Finally, we would love to extend our diversity result (Corollary 4) to three or more genes. Unfortunately, in this more general case the equations become multilinear, and such equations are very hard both to solve and to argue about. Our simulations, however, show that large equilibria exist in the multi-gene setting as well.

Our discussion has focused on the evolution of a fixed set of alleles; that is, we have not discussed mutations. Mutations are, of course, paramount in evolution, as they are the source of genetic diversity, and we believe that introducing mutations to the present analysis is an important research direction. Here we focus on the selection process, which is rigorously shown to exploit the diversity created by mutations to enhance expected fitness, while at the same time also maintaining this diversity.

We can now note a simple yet important point. Since MWUA only operates in the presence of sex, the triple connection uncovered in this paper is informative for the “queen of problems in evolutionary biology,” namely the role of sex in evolution [13, 10]. The notion that the role of sex is the maintenance of diversity has been critiqued [111], since sex does not always increase diversity, and diversity is not always favorable. The MWUA connection sheds new light on the debate, since sex is shown to lead to a tradeoff between increasing entropy and increasing (cumulative) fitness.

The connection between the three fields, evolution, game theory, and learning algorithms, described here was not accessible to the founders of the modern synthesis, and we hope that it expands the mathematical tracks that can be traveled in evolution theory.

The recently established bridge between evolutionary biology and theoretical computer science has proven productive within a short time-span. While evolutionary theory was founded on the mathematical tools of physics and statistics, it actually shares much with the younger field of theoretical computer science, namely an interest in complexity and performance, or “fitness.” The discovery of mixability led to a pursuit of a mixability-based maximization principle for evolution [25], which then led to the finding of a maximization principle for evolution in the realm of weak selection [25]. As Barton et al. [11] noted, “the close analogy between MWUA and population

genetics is a first step” in the process of a fruitful transfer of knowledge between these two fields, and indeed, this analogy allowed us to recognize the applicability of the MWUA no-regret theorem in population genetics. There has already been tremendous progress in applying these methods to population genetics, with Ioannis Panageas and co-authors resolving an open conjecture dating from the 1970’s using these tools [94] (showing that haploid populations will end up not having any diversity asymptotically). Using the learning perspective that we have advanced for population genetics, others have followed-up and considered how learning algorithms could be the basis for the study of genetics in a broader sense [143]. The full impact and interpretation of this theorem is yet to be seen.

Some additional things we would like to work on in the future include:

- Is the diploid population PG dynamics (infinite population selection with recombination) also the MWUA in a suitable game? We also would like to see a similar result for the finite population case....

In particular, note that in the Diploid case there is a form of MWU that describes the dynamics for the marginal frequencies of the alleles. However, the quality function $-l_i$ ends up being something other than the expected value for the payoff of a pure strategy in a game, though it is the product of the expected payoff with the probability of choosing that pure strategy. For the finite population case, first one would need to extend Nagylaki’s theorem to the finite population case. It is currently only for an infinite population setting. An extension of Nagylaki’s theorem to a finite population setting would require serious work in Markov Chain theory, as the typical model for finite population evolution (Wright-Fisher or the Moran Process model) would be a Markov Chain. In addition to the extension of Nagylaki’s theorem, one would have to extend the connection between MWU and Game theory too. If one observes the form of the Moran process, it looks very much like the MWU. Could that connection be used to show an analogous result for a coordination game assuming weak selection?

- A more general theory for MWU on coordination games, for instance providing convergence times for different coordination games. Such would give us insights into the kinds of fitness landscapes that lead to fast convergence.

8.2 Evolution of Animal Traits

Besides the application of the MWUA to Population Genetics, we also showed that the MWU method could be used for modeling the Evolution of Animal Personality and Tool-innovation.

The new model of tool-innovation is very rich, and has many different unexplored issues. Most of these have to do with the complexities of analysis relating to random sampling of polynomially-valued functions. We would like to see whether we can expand the elementary analysis based on the mean-value theorem to non-uniform distributions over subsets of the real line, and thereby give a more interesting characterization of when tool-innovation would evolve and when it wouldn't. To do this, one would need to give a new set of inequalities for the expected value of the difference between a sampled point's value and the maximum value.

We also want to analyze more properties of the Tool-innovation model for quality functions that are multi-variate polynomials. We have so far considered only univariate polynomials, but for complex tool-innovation tasks one would expect a satisfiability problem to be a more useful framework, and that necessitates the use of multivariate polynomials. However the current analytic tools all assume ordinary univariate polynomials.... Perhaps one could use a suitable multi-variate generalization of the mean-value theorem?

For the animal personality model, could we apply some of the ideas there developed to give constraints on when no animal personality at all would evolve? How rare is such an eventuality, or is it actually quite common in the model that animal personality will emerge? Some of the counter-examples to animal personality emerging seem to come from very predictable environments. How many of these are there, combinatorially speaking?

8.3 Over-arching themes

The idea of Multiplicative Weight Updates being so useful as a way of generating models for evolutionary biology brings up a few general questions:

- Are there other multi-purpose online learning algorithms that are generally useful for evolutionary modeling, or is MWUA in some sense special?
- Could the versatility of the MWUA approach to modeling biological dynamics be related in some formal way to the success of approaches in Neuroscience, Psychology, and Machine Learning that use Bayesian probability updating as a basic modeling framework?
- Are there other algorithmic tools or complexity-theoretic tools that could be useful for population genetics or evolutionary theory?
- Is there something special about algorithmic or computational kinds of models in biology, or is it a coincidence that MWUA is so useful in this field? Could there be some kind of base model, for instance, that allows us to make more structured models using algorithmic theory than simple bean-bag genetics would allow?

Importantly, the tools that computational complexity, discrete analysis (of boolean functions), and other similar tools from theoretical computer science study how small changes in problem representation or specification could lead to dramatic changes in encoded functions. Such analysis could be related to small changes in “hardware” leading to dramatic changes in “software.” In a similar way, one could use the same tools to analyze rigorously and predict large changes in phenotype based on small changes in genotype.

In fact, the broader questions of how major transitions in evolution happen, or major changes occurs in phenotype. Some of the work in this thesis was focused on this question in more particular domains, applying computational thinking to various population genetics and ecology questions involving major transitions (evolution of sex, and animal cognition among others). Now a new direction would be to find out

how biological information-processing mechanisms could arise and change radically and irreversibly with small genetic changes. Could it be perhaps that using something like “invariance principles” one could show that major transitions in evolution occur by virtue of something like the central limit theorem applied to the effects on the phenotype of different subsets of alleles?

Recall one of the kinds of models in Evolutionary biology was “arrival of the fittest”: Would the same varieties recur if evolution were replayed again from the same initial condition? An interesting sub-theme of this question of arrival of the fittest is to analyze the “bias” and “variance” of evolution. By that I mean, the expected divergence from the current trajectory, and the variance of the divergence from the current trajectory. To do this, one would try to analyze various population genetics models from the perspective of learning theory, with evolutionary dynamics being considered learning algorithms. The work outlined in this thesis has explored some of these themes, but I would like to continue to see in what way we could for instance apply statistical learning theory or online learning to understand these questions.

References

- [1] Shahal Abbo, Dan Shtienberg, Judith Lichtenzveig, Simcha Lev-Yadun, and Avi Gopher. The chickpea, summer cropping, and a new model for pulse domestication in the ancient near east. *The Quarterly Review of Biology*, 78(4):435–448, 2003.
- [2] John Alcock. The evolution of the use of tools by feeding animals. *Evolution*, pages 464–473, 1972.
- [3] David Aldous and Umesh Vazirani. “go with the winners” algorithms. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 492–501. IEEE, 1994.
- [4] Martin Anthony and Peter L Bartlett. *Neural network learning: Theoretical foundations*. Cambridge University Press, 2009.
- [5] Aristotle. *De interpretatione*.
- [6] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.
- [7] Sanjeev Arora, Yuval Rabani, and Umesh Vazirani. Simulating quadratic dynamical systems is pspace-complete (preliminary version). In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 459–467. ACM, 1994.
- [8] Sancti Augustini. *Confessiones*.
- [9] Nicholas H. Barton, Derek E. G. Briggs, Jonathan A. Eisen, David B. Goldstein, and Nipam H. Patel. *Evolution*. Cold Spring Harbor Laboratory Press, 2007.
- [10] Nicholas H Barton and Brian Charlesworth. Why sex and recombination? *Science*, 281(5385):1986–1990, 1998.
- [11] Nicholas H Barton, Sebastian Novak, and Tiago Paixão. Diverse forms of selection in evolution and computer science. *Proceedings of the National Academy of Sciences*, 111(29):10398–10399, 2014.
- [12] Alison M Bell and Andrew Sih. Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecology Letters*, 10(9):828–834, 2007.
- [13] G Bell. *The Masterpiece of Nature: the Evolution and Genetics of Sexuality*. University of California Press, Berkeley, 1982.

- [14] CT Bergstrom and LA Dugatkin. *Evolution*. WW Norton and Co, New York, 2011.
- [15] Christopher D Bird and Nathan J Emery. Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proceedings of the National Academy of Sciences*, 106(25):10370–10375, 2009.
- [16] Peter A Biro and Judy A Stamps. Are animal personality traits linked to life-history productivity? *Trends in Ecology & Evolution*, 23(7):361–368, 2008.
- [17] Lucas A Bluff, Alex Kacelnik, and Christian Rutz. Vocal culture in New Caledonian crows *Corvus moneduloides*. *Biological journal of the Linnean Society*, 101(4):767–776, 2010.
- [18] Peter Borwein and Tamás Erdélyi. *Polynomials and polynomial inequalities*, volume 161. Springer Science & Business Media, 2012.
- [19] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [20] R. Bürger. Some mathematical models in evolutionary genetics. *The Mathematics of Darwin's Legacy*, pages 67–89, 2011.
- [21] Reinhard Bürger. *The mathematical theory of selection, recombination, and mutation*, volume 228. Wiley Chichester, 2000.
- [22] John Curtis Burkhalter, Nina H Fefferman, and Julie L Lockwood. The impact of personality on the success of prospecting behavior in changing landscapes. *Current Zoology*, 61(3):557–568, 2015.
- [23] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- [24] Erick Chastain, Adi Livnat, Christos Papadimitriou, and Umesh Vazirani. Multiplicative updates in coordination games and the theory of evolution. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 57–58. ACM, 2013.
- [25] Erick Chastain, Adi Livnat, Christos Papadimitriou, and Umesh Vazirani. Algorithms, games, and evolution. *PNAS*, 111(29), June 2014.
- [26] Bernard Chazelle. Analytical tools for natural algorithms. In *ICS*, pages 32–41, 2010.
- [27] Bernard Chazelle. The convergence of bird flocking. *Journal of the ACM (JACM)*, 61(4):21, 2014.
- [28] Freddy Bugge Christiansen. *Population genetics of multiple loci*. John Wiley & Sons Ltd, 1999.
- [29] Thomas M Cover. Universal portfolios. *Mathematical finance*, 1(1):1–29, 1991.
- [30] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

- [31] JF Crow. Genetic loads and the cost of natural selection. In *Mathematical topics in population genetics*, pages 128–177. Springer, 1970.
- [32] J.F. Crow and M. Kimura. Evolution in sexual and asexual populations. *American Naturalist*, pages 439–450, 1965.
- [33] Sasha RX Dall, Alasdair I Houston, and John M McNamara. The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters*, 7(8):734–739, 2004.
- [34] J. Darling. *Song*. In *Encyclopedia of marine mammals*. Academic Press, 2009.
- [35] Charles Darwin. On the origin of the species by natural selection. 1859.
- [36] Richard Dawkins. *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company, 1986.
- [37] Derrick De Kerckhove, Robert L McLaughlin, and David LG Noakes. Ecological mechanisms favouring behavioural diversification in the absence of morphological diversification: a theoretical examination using brook charr (*Salvelinus fontinalis*). *Journal of Animal Ecology*, 75(2):506–517, 2006.
- [38] Daniel Dennett. *Darwin’s Dangerous Idea*. Simon & Schuster, 1995.
- [39] Niels J Dingemanse and Max Wolf. Recent models for adaptive personality differences: a review. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1560):3947–3958, 2010.
- [40] Theodosius Dobzhansky. A review of some fundamental concepts and problems of population genetics. In *Cold Spring Harbor Symp Quant Biol*, volume 20, pages 1–15. Cold Spring Harbor Laboratory Press, 1955.
- [41] Theodosius Dobzhansky. Nothing in biology makes sense except in the light of evolution. *The American biology teacher*, 75(2):87–91, 2013.
- [42] DN Duvick and Kenneth G Cassman. Post-green revolution trends in yield potential of temperate maize in the north-central united states. *Crop science*, 39(6):1622–1630, 1999.
- [43] Manfred Eigen. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523, 1971.
- [44] J Michael Fay and Richard W Carroll. Chimpanzee tool use for honey and termite extraction in central africa. *American Journal of Primatology*, 34(4):309–317, 1994.
- [45] Marcus W Feldman, Sarah P Otto, and Freddy B Christiansen. Population genetic perspectives on the evolution of recombination. *Annual review of genetics*, 30(1):261–295, 1996.
- [46] PC Fishburn and NJA Sloane. The solution to Berlekamp’s switching game. *Discrete mathematics*, 74(3):263–290, 1989.
- [47] Ronald A Fisher. *The genetic theory of natural selection*. Dover, 1958.

- [48] Walter Fontana and Leo W Buss. “The arrival of the fittest”: Toward a theory of biological organization. *Bulletin of Mathematical Biology*, 56(1):1–64, 1994.
- [49] Paul C Foster, Nathan J Mlot, Angela Lin, and David L Hu. Fire ants actively control spacing and orientation within self-assemblages. *Journal of Experimental Biology*, 217(12):2089–2100, 2014.
- [50] Pope Francis. *Laudato si: On care for our common home*. Our Sunday Visitor, 2015.
- [51] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer, 1995.
- [52] Ellen C Garland, Jason Gedamke, Melinda L Rekdahl, Michael J Noad, Claire Garrigue, and Nick Gales. Humpback whale song on the southern ocean feeding grounds: implications for cultural transmission. *PloS one*, 8(11):e79422, 2013.
- [53] Andrew Garrard. Charting the emergence of cereal and pulse domestication in south-west Asia. *Environmental Archaeology*, 1999.
- [54] Johann Wolfgang von Goethe. An intermaxillary bone is present in the upper jaw of man as well as in animals. *Goethe, Scientific Studies*, pages 115–16, 1786.
- [55] Ernst Haeckel. *Natürliche Schöpfungsgeschichte: Gemeinverständliche wissenschaftliche Vorträge über die Entwicklungslehre im Allgemeinen und diejenige von Darwin, Göthe und Lamarck im Besonderen, über die Anwendung derselben auf den Ursprung des Menschen und andern damit zusammenhängende Gründfragen der Natur-Wissenschaft. Mit Tafeln, Holzschnitten, systematischen und genealogischen Tabellen*. Gg. Reimer, 1868.
- [56] James Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.
- [57] Yan Hao, Dieter Armbruster, Lee Cronk, and C Athena Aktipis. Need-based transfers on a network: a model of risk-pooling in ecologically volatile environments. *Evolution and Human Behavior*, 36(4):265–273, 2015.
- [58] Yan Hao, Dieter Armbruster, and Marc-Thorsten Hütt. Node survival in networks under correlated attacks. *PloS one*, 10(5):e0125467, 2015.
- [59] Daniel L Hartl and Vitezslav Orel. What did Gregor Mendel think he discovered? *Genetics*, 131(2):245, 1992.
- [60] John H Holland. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press, 1975.
- [61] Jennifer C Holzhaider, Gavin R Hunt, and Russell D Gray. The development of pandanus tool manufacture in wild New Caledonian crows. *Behaviour*, 147(5):553–586, 2010.

- [62] Gavin R Hunt and Russell D Gray. Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1517):867–874, 2003.
- [63] Jennifer M Jandt, Sarah Bengston, Noa Pinter-Wollman, Jonathan N Pruitt, Nigel E Raine, Anna Dornhaus, and Andrew Sih. Behavioural syndromes and social insects: personality at multiple levels. *Biological Reviews*, 89(1):48–67, 2014.
- [64] Peter R Jennings and Jose de Jesus. Studies on competition in rice I. competition in mixtures of varieties. *Evolution*, 22(1):119–124, 1968.
- [65] D. Johnson, C. Papadimitriou, and M. Yannakakis. How easy is local search? *Journal of computer and system sciences*, 37(1):79–100, 1988.
- [66] J Chadwick Johnson and Andrew Sih. Precopulatory sexual cannibalism in fishing spiders (*Dolomedes triton*): a role for behavioral syndromes. *Behavioral Ecology and Sociobiology*, 58(4):390–396, 2005.
- [67] Brendan Juba. *Universal semantic communication*. Springer Science & Business Media, 2011.
- [68] Brendan Juba and Madhu Sudan. Efficient semantic communication via compatible beliefs. 2011.
- [69] Brendan Juba and Santosh Vempala. Semantic communication for simple goals is equivalent to on-line learning. In *Algorithmic Learning Theory*, pages 277–291. Springer, 2011.
- [70] Hans Kalmus. The scholastic origins of Mendel’s concepts. *History of science*, 21(1):61–83, 1983.
- [71] John L Kelly Jr. A new interpretation of information rate. *Information Theory, IRE Transactions on*, 2(3):185–189, 1956.
- [72] Ben Kenward, Christian Rutz, Alex AS Weir, and Alex Kacelnik. Development of tool use in New Caledonian crows: inherited action patterns and social influences. *Animal Behaviour*, 72(6):1329–1343, 2006.
- [73] Ben Kenward, Christian Schloegl, Christian Rutz, Alexander AS Weir, Thomas Bugnyar, and Alex Kacelnik. On the evolutionary and ontogenetic origins of tool-oriented behaviour in New Caledonian crows (*Corvus moneduloides*). *Biological Journal of the Linnean Society*, 102(4):870–877, 2011.
- [74] M Kimura. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, 1985.
- [75] Motoo Kimura. Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626, 1968.
- [76] David C Krakauer and Joshua B Plotkin. Redundancy, antiredundancy, and the robustness of genomes. *Proceedings of the National Academy of Sciences*, 99(3):1405–1409, 2002.

- [77] Edo Kussell and Stanislas Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–8, 2005.
- [78] Michael Lachmann and Carl T Bergstrom. The disadvantage of combinatorial communication. *Proceedings of the Royal Society of London B: Biological Sciences*, 271(1555):2337–2343, 2004.
- [79] Andre Levchenko and Ilya Nemenman. Cellular noise and information transmission. *Current opinion in biotechnology*, 28:156–164, 2014.
- [80] Richard Levins and Richard C Lewontin. *The dialectical biologist*. Harvard University Press, 1987.
- [81] RC Lewontin and JL Hubby. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, 54:595–609, 1966.
- [82] Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine learning*, 2(4):285–318, 1988.
- [83] A. Livnat, C. Papadimitriou, J. Dushoff, and M.W. Feldman. A mixability theory for the role of sex in evolution. *Proceedings of the National Academy of Sciences*, 105(50):19803–19808, 2008.
- [84] Adi Livnat, Christos Papadimitriou, Jonathan Dushoff, and Marcus W Feldman. A mixability theory for the role of sex in evolution. *Proceedings of the National Academy of Sciences*, 105(50):19803–19808, 2008.
- [85] Adi Livnat, Christos Papadimitriou, and Marcus W Feldman. An analytical contrast between fitness maximization and selection for mixability. *Journal of theoretical biology*, 273(1):232–234, 2011.
- [86] Adi Livnat, Christos Papadimitriou, Nicholas Pippenger, and Marcus W Feldman. Sex, mixability, and modularity. *Proceedings of the National Academy of Sciences*, 107(4):1452–1457, 2010.
- [87] Adi Livnat, Christos Papadimitriou, Aviad Rubinstein, Gregory Valiant, and Andrew Wan. Satisfiability and evolution. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 524–530. IEEE, 2014.
- [88] Salvador E Luria and Max Delbrück. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28(6):491, 1943.
- [89] Saunders Mac Lane. *Categories for the working mathematician*, volume 5. Springer Science & Business Media, 1978.
- [90] Exu Anton Mates, Robin R Tarter, James C Ha, Anne B Clark, and Kevin J McGowan. Acoustic profiling in a complexly social species, the American crow: caws encode information on caller sex, identity and behavioural context. *Bioacoustics*, 24(1):63–80, 2015.
- [91] J Maynard-Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.

- [92] E Mayr. *Animal Species and Evolution*. Belknap Press, 1963.
- [93] Daniel J McDonald. Generalization error bounds for time series. *Carnegie Mellon University Doctoral Dissertation for the Department of Statistics*, 2012.
- [94] Ruta Mehta, Ioannis Panageas, and Georgios Piliouras. Natural selection as an inhibitor of genetic diversity: Multiplicative weights updates algorithm and a conjecture of haploid genetics [working paper abstract]. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, pages 73–73. ACM, 2015.
- [95] Reshef Meir and David Parkes. On sex, evolution, and the multiplicative weights update algorithm. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 929–937. International Foundation for Autonomous Agents and Multiagent Systems, 2015.
- [96] Dusan Misevic, Charles Ofria, and Richard E Lenski. Sexual reproduction reshapes the genetic architecture of digital organisms. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1585):457–464, 2006.
- [97] George Molnar and Stephen Mumford. *Powers: A study in metaphysics*. Oxford University Press on Demand, 2003.
- [98] Pierre-Olivier Montiglio, Tina W Wey, Ann T Chang, Sean Fogarty, and Andrew Sih. Multiple mating reveals complex patterns of assortative mating by personality and body size. *Journal of Animal Ecology*, 85(1):125–135, 2016.
- [99] Lynne M Mullen and Hopi E Hoekstra. Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution*, 62(7):1555–1570, 2008.
- [100] T. Nagylaki. The evolution of multilocus systems under weak selection. *Genetics*, 134(2):627–647, 1993.
- [101] T. Nagylaki, J. Hofbauer, and P. Brunovský. Convergence of multilocus systems under weak epistasis or weak selection. *Journal of mathematical biology*, 38(2):103–133, 1999.
- [102] Thomas Nagylaki. *Introduction to theoretical population genetics*, volume 142. Springer-Verlag Berlin, 1992.
- [103] Thomas Nagylaki, Josef Hofbauer, and Pavol Brunovský. Convergence of multilocus systems under weak epistasis or weak selection. *J Math Biol*, 38(2):103–133, 1999.
- [104] Masatoshi Nei. Selectionism and neutralism in molecular evolution. *Mol Biol Evol*, 22(12):2318–2342, 2005.
- [105] Michael J Noad, Douglas H Cato, MM Bryden, Micheline-N Jenner, and K Curt S Jenner. Cultural revolution in whale songs. *Nature*, 408(6812):537–537, 2000.
- [106] David S Oderberg. *Real essentialism*. Routledge, 2007.
- [107] Charles Kay Ogden, Ivor Armstrong Richards, Bronislaw Malinowski, and Francis Graham Crookshank. *The meaning of meaning*. Kegan Paul London, 1923.

- [108] Saint Thomas Aquinas O.P. *Commentary on Aristotle's Physics III-10*. Yale University Press, 1963. Translated by Richard J. Blackwell, Richard J. Spath and W. Edmund Thirlkel.
- [109] Lorenzo Orecchia and M Mahoney. Implementing regularization implicitly via approximate eigenvector computation. *Proc 28th Intl Conf Machine Learning*, pages 121–128, 2011.
- [110] R Ortiz-Monasterio, KD Sayre, S Rajaram, M McMahon, et al. Genetic progress in wheat yield and nitrogen use efficiency under four nitrogen rates. *Crop Science*, 37(3):898–904, 1997.
- [111] Sarah P Otto. The evolutionary enigma of sex. *Am Nat*, 174(S1):S1–S14, 2009.
- [112] Katharine Payne and Roger Payne. Large scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift für Tierpsychologie*, 68(2):89–114, 1985.
- [113] Daniel J Povinelli, James E Reaux, Laura A Theall, and Steve Giambrone. The question of tool modification. *Folk physics for apes*. Oxford University Press, Oxford, UK, pages 271–296, 2000.
- [114] Jonathan N Pruitt. Animal personality: Ecology, behavior, and evolution. *Psychological science*, 8:69–75, 2014.
- [115] Richard O Prum and Alan H Brush. The evolutionary origin and diversification of feathers. *The Quarterly review of biology*, 77(3):261–295, 2002.
- [116] Rami Pugatch, Naama Barkai, and Tsvi Tlusty. Asymptotic cellular growth rate as the effective information utilization rate. *arXiv preprint arXiv:1308.0623*, 2013.
- [117] Adam R Reddon. Parental effects on animal personality. *Behavioral Ecology*, page 210, 2011.
- [118] Chris R Reid, Matthew J Lutz, Scott Powell, Albert B Kao, Iain D Couzin, and Simon Garnier. Army ants dynamically adjust living bridges in response to a cost-benefit trade-off. *Proceedings of the National Academy of Sciences*, 112(49):15113–15118, 2015.
- [119] Olivier Rivoire. Information in models of evolutionary dynamics. *arXiv*, (21):1–22, 2015.
- [120] Olivier Rivoire and Stanislas Leibler. The Value of Information for Populations in Varying Environments. *Journal of Statistical Physics*, 142(6):1124–1166, March 2011.
- [121] Olivier Rivoire and Stanislas Leibler. A model for the generation and transmission of variations in evolution. *PNAS*, 111(15):E1940–9, December 2014.
- [122] Areejit Samal, João F Matias Rodrigues, Jürgen Jost, Olivier C Martin, and Andreas Wagner. Genotype networks in metabolic reaction spaces. *BMC systems biology*, 4(1):1, 2010.

- [123] Peter Schuster and Karl Sigmund. Replicator dynamics. *Journal of theoretical biology*, 100(3):533–538, 1983.
- [124] Claude E Shannon and Warren Weaver. *The mathematical theory of communication*. University of Illinois press, 2015.
- [125] Karl Sigmund. A maximum principle for frequency dependent selection. *Mathematical Biosciences*, 84:189–195, 1987.
- [126] Andrew Sih, Ann T Chang, and Tina W Wey. Effects of behavioural type, social skill and the social environment on male mating success in water striders. *Animal Behaviour*, 94:9–17, 2014.
- [127] Andrew Sih, Julien Cote, Mara Evans, Sean Fogarty, and Jonathan Pruitt. Ecological implications of behavioural syndromes. *Ecology Letters*, 15(3):278–289, 2012.
- [128] Andrew Sih, Kimberley J Mathot, Maria Moiron, Pierre-Olivier Montiglio, Max Wolf, and Niels J Dingemans. Animal personality and state-behaviour feedbacks: a review and guide for empiricists. *Trends in Ecology & Evolution*, 30(1):50–60, 2015.
- [129] Andrew Sih and Jason V Watters. The mix matters: behavioural types and group dynamics in water striders. *Behaviour*, 142(9-10):1417–1431, 2005.
- [130] John Maynard Smith. The evolution of animal intelligence. In *Did Darwin Get It Right?*, pages 192–201. Springer Science Business Media, 1988.
- [131] David I Spivak. *Category Theory for the Sciences*. MIT Press, 2014.
- [132] E. Szathmari and J. Maynard Smith. The major transitions in evolution. *Nature*, 374:227–232, 1995.
- [133] Jack W Szostak. Functional information: Molecular messages. *Nature*, 423(6941):689–689, 2003.
- [134] Gašper Tkačik and William Bialek. Information processing in living systems. *arXiv preprint arXiv:1412.8752*, 2014.
- [135] Larissa Trompf and Culum Brown. Personality affects learning and trade-offs between private and social information in guppies, *Poecilia reticulata*. *Animal Behaviour*, 88:99–106, 2014.
- [136] Koji Tsuda, Gunnar Rätsch, and Manfred K Warmuth. Matrix exponentiated gradient updates for on-line learning and bregman projection. In *J Mach Learn Res*, pages 995–1018, 2005.
- [137] Leslie G Valiant. Evolvability. *J ACM*, 56(1):3, 2009.
- [138] Alexander Von Humboldt. *Views of Nature: Or Contemplations on the Sublime Phenomena of Creation. With Scientific Illustrations*. G. Bell, 1850.
- [139] John Von Neumann, Arthur W Burks, et al. Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, 5(1):3–14, 1966.

- [140] Andreas Wagner. *Arrival of the Fittest: Solving Evolution's Greatest Puzzle*. Penguin, 2014.
- [141] Gunter P Wagner and Lee Altenberg. Perspective: complex adaptations and the evolution of evolvability. *Evolution*, pages 967–976, 1996.
- [142] Günter P Wagner, Mihaela Pavlicev, and James M Cheverud. The road to modularity. *Nature Reviews Genetics*, 8(12):921–931, 2007.
- [143] Richard A. Watson and Eors Szathmary. How can evolution learn? *Trends in Ecology & Evolution*, 2015.
- [144] J W Weibull. *Evolutionary Game Theory*. MIT Press, 1995.
- [145] Edward D Weinberger. A theory of pragmatic information and its application to the quasi-species model of biological evolution. *Biosystems*, 66(3):105–119, 2002.
- [146] Alex AS Weir, Jackie Chappell, and Alex Kacelnik. Shaping of hooks in New Caledonian crows. *Science*, 297(5583):981–981, 2002.
- [147] Alex AS Weir and Alex Kacelnik. A New Caledonian crow (*Corvus moneduloides*) creatively re-designs tools by bending or unbending aluminium strips. *Animal cognition*, 9(4):317–334, 2006.
- [148] Tina W Wey, Ann T Chang, Sean Fogarty, and Andrew Sih. Personalities and presence of hyperaggressive males influence male mating exclusivity and effective mating in stream water striders. *Behavioral Ecology and Sociobiology*, 69(1):27–37, 2015.
- [149] Max Wolf, G Sander Van Doorn, and Franz J Weissing. Evolutionary emergence of responsive and unresponsive personalities. *Proceedings of the National Academy of Sciences*, 105(41):15825–15830, 2008.
- [150] Colin M Wright, C Tate Holbrook, and Jonathan N Pruitt. Animal personality aligns task specialization and task proficiency in a spider society. *Proceedings of the National Academy of Sciences*, 111(26):9533–9537, 2014.
- [151] Jessica L Yorzinski, Sandra L Vehrencamp, Kevin J McGowan, and Anne B Clark. The inflected alarm caw of the American crow: differences in acoustic structure among individuals and sexes. *The Condor*, 108(3):518–529, 2006.
- [152] Daniel Zohary, Maria Hopf, and Ehud Weiss. *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. Oxford University Press on Demand, 2000.