

Adaptation in a Changing Environment:
Phenotypic Diversity in Response to
Environmental Uncertainty and Information

Matina C. Donaldson-Matasci

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2008

Program Authorized to Offer Degree: Biology

University of Washington

Abstract

Adaptation in a Changing Environment:
Phenotypic Diversity in Response to
Environmental Uncertainty and Information

Matina C. Donaldson-Matasci

Chair of the Supervisory Committee:
Professor Carl T. Bergstrom
Department of Biology

Organisms living in a changing environment must be adapted for a wide range of conditions. However, it is often difficult to be well-equipped for all possible situations at once. One way that individuals can overcome fitness trade-offs is by responding to change over the course of their lifetimes. Another is developmental plasticity, in which a single genotype gives rise to multiple phenotypes suited for different environmental conditions. Both can produce remarkable diversity in morphology and behavior within a species. How do organisms evolve to produce and direct that diversity?

In this dissertation, I model the evolution of plasticity in three contexts. First, some animals that can respond to unexpected events like predator attacks also develop communication systems about those events. The way that signals become linked to meanings, and the power of the communication system to discriminate among different events, depends on fitness trade-offs. Second, some organisms adapt to unpredictable environments by randomly producing a mixture of phenotypes suited for different conditions, effectively hedging bets against catastrophic events. I show how the optimal amount of diversity produced depends on the amount of uncertainty in the environment. Third, individuals can sometimes use cues gathered during development to help predict the future environment, and better direct their developmental trajectory. I show how the value of this predictive information depends on the way it is distributed among individuals within and between generations. Taken together, the three chapters demonstrate several ways that plastic strategies come to reflect the patterns of change and predictability in the environment to which they are adapted.

TABLE OF CONTENTS

	Page
List of Figures	iii
List of Tables	iv
Chapter 1: Introduction	1
1.1 Bet-hedging as an adaptation to environmental uncertainty	2
1.2 Predictive plasticity as an adaptation to environmental information	3
1.3 The impact of environmental variation depends on fitness trade-offs	4
1.4 An outline of the dissertation	5
Chapter 2: Behavioral flexibility and communication: How fitness trade-offs shape animal signalling systems	7
2.1 Introduction	7
2.2 A model for the evolution of communication	9
2.3 Evolutionary stability of communication systems	12
2.4 Conclusions	21
Chapter 3: Developmental plasticity and environmental risk: How phenotypic diversity reflects the amount of risk	24
3.1 Introduction	24
3.2 Background: Levins's fitness sets and evolution in an uncertain environment	25
3.3 A model of the evolution of generalists	28
3.4 Generality of the model	42
3.5 Discussion	44
Chapter 4: Developmental plasticity and environmental information: How the value of information depends on its structure	51
4.1 Introduction	51
4.2 A model of developmental plasticity in response to unreliable cues	52
4.3 Optimal developmental strategies and the value of a cue	59

4.4 Discussion 73

LIST OF FIGURES

Figure Number	Page
2.1 Sample payoff matrices for the Hurford-Nowak model and our model.	11
2.2 An illustration of signal pooling.	14
2.3 Different signal poolings can yield different ESSs for the same system.	16
2.4 An example of a communication system with an ES set consisting of a single signalling strategy and multiple receiving strategies.	17
2.5 Meaning can cycle continually when there are more signals than states or actions. .	19
3.1 Levins's fitness set and adaptive function.	27
3.2 The fitness set for an example of frog metamorphosis.	31
3.3 The extended fitness set for an example of frog metamorphosis.	32
3.4 An optimal bet-hedging strategy matches its specialization in each environment to the frequency of that environment.	35
3.5 Individual-level uncertainty reduces selection for phenotypic diversification.	39
3.6 A fixed cost to plasticity reduces the range of environmental uncertainty in which bet-hedging is optimal.	41
4.1 Types of developmental strategies.	56
4.2 Optimal developmental strategies for a cue with population-level error.	63
4.3 The value of a cue with population-level error is the information it provides about the environment.	64
4.4 Optimal developmental strategies for a cue with individual-level error.	67
4.5 The value of a cue with individual-level error may exceed the information it contains about the environment.	68
4.6 Optimal developmental strategies for a cue with error at both levels.	71
4.7 The value of a cue is the information it provides to the population as a whole. . . .	72

LIST OF TABLES

Table Number	Page
4.1 Summary of mathematical notation.	54

ACKNOWLEDGMENTS

The author wishes to thank first and foremost Carl Bergstrom and Michael Lachmann, with whom all of this work was done. Without their guidance and collaborative spirit, none of this would have been possible. The following people read some or all of the dissertation, and made helpful comments and suggestions: Chris Templeton, Jelle Zuidema, Franjo Weissing, Lauri Oksanen, Ben Kerr, Joe Felsenstein, and Ray Huey. Saskia Werners and Dean Foster made useful suggestions for Chapter 3. Nils Bertschinger and Eckehard Olbrich provided helpful discussion of Chapter 4. Arthur Robson suggested the use of de Finetti's theorem in the Appendix to Chapter 4. Much of this work was done under the support of an NSF Graduate Research Fellowship; further funding came from the UW Department of Biology, the ARCS Foundation, and the Max Planck Society. Finally, the author expresses sincere appreciation to Naim Matasci, whose intellectual and moral support was critical throughout.

DEDICATION

to my grandfather, who started this family tradition.

Chapter 1

INTRODUCTION

Every species experiences environmental variation at some level. Dispersing seeds may encounter different soil types, moisture levels, and access to sunlight, depending on exactly where they happen to land. From one year to the next, the ideal conditions for germination—the right temperature and sufficient rainfall—may come at different times, or not at all. Many seed-eating birds eat a wide variety of seeds that differ size and hardness. The proportions of each seed type can vary greatly depending on habitat type, the season, how dry the year has been, and how much competition there is for each seed type. These birds in turn may be subject to predation by different kinds of predators with different hunting techniques. Whether any individual bird is attacked is partly a matter of chance, but also depends on the density of predators and prey in the region.

All this variation in environmental conditions can have an enormous impact on the fitness of different life history traits, morphologies, and behaviors. It makes adaptation to any one set of circumstances difficult, because phenotypes that are favored in some conditions may be selected against in others. One way to avoid the problem of adaptation to multiple environments is for individuals to actively choose a habitat. For example, a bird may be able to selectively forage in areas where its preferred kind of seed is particularly abundant. In this way, the range of food sources utilized can be greatly reduced, making it possible to specialize effectively on just one type (Templeton and Rothman, 1981). On the other hand, it will often not be possible to completely eliminate all environmental variation, and in many cases it may even be counterproductive (Wilson and Yoshimura, 1994). In this dissertation, I will not consider the ecological niche of the species as an evolvable trait. Rather, I focus on strategies that allow a species to tolerate the entire suite of environmental conditions it may encounter.

One way for a species to have a broad environmental tolerance is for each individual to be a generalist. This is sometimes accomplished via a fixed generalist phenotype, which performs

reasonably well under a wide variety of circumstances. A flexible generalist, which can respond to changing conditions by changing its behavior or morphology, may be able to do even better. On the other hand, if individuals are specialists, flourishing only under a narrow set of conditions, the species as a whole may nonetheless be quite tolerant of change as long as there is substantial variation between individuals. This variation may stem from genetic differences, from differences in the developmental environment, or it may be entirely random.

In this dissertation, I examine the adaptive value of phenotypic diversity, both within and between individuals, in the context of environmental variation. How does the amount of uncertainty in the environment, and the kind of information available, affect the types of strategies used to generate phenotypic diversity? To what extent do those strategies reflect patterns of environmental change and predictability? I draw upon several well-developed themes in the theoretical biology literature: bet-hedging, predictive plasticity, and fitness trade-offs.

1.1 Bet-hedging as an adaptation to environmental uncertainty

Environmental conditions may vary unpredictably over an enormous range of scales in both time and space. The way that individuals experience these variations depends on how long they live, and how often and how far they move within a lifetime (Levins, 1968). The way that the population as a whole is affected depends on how similar the experiences of different individuals are. Although a contrast is often drawn between spatial and temporal variation, the most important distinction lies in the way that individual variation in fitness contributes to variation in a lineage's growth rate (Frank and Slatkin, 1990). The growth of a lineage in one generation depends on the sum of fitnesses of all individuals, but growth over multiple generations depends on the product of growth in each generation. For this reason, the evolutionary impact of *individual-level risk* (environmental variation between individuals within a generation) is qualitatively different from the impact of *population-level risk* (environmental variation between generations.)

The first models to consider the evolutionary consequences of population-level environmental uncertainty were specifically concerned with the maintenance of genetic polymorphism (Dempster, 1955; Haldane and Jayakar, 1962). In this arena risk structure does not seem to have a dramatic impact—for variation at either level, although a polymorphic population may maximize the popula-

tion growth rate, polymorphism cannot be maintained except through some form of overdominance (Gillespie, 1973). A genetic polymorphism therefore cannot generally be considered an adaptation to environmental variation *per se* (Felsenstein 1976; but see Leimar 2005.) However, if a single genotype can generate random phenotypic polymorphism, in proportions that optimize the growth rate of the lineage over the long term, this genotype can be selected for. Adaptive phenotypic variation of this sort is known as *bet-hedging* (Seeger and Brockmann, 1987) or *adaptive coin-flipping* (Cooper and Kaplan, 1982).

The first model to consider random phenotypic plasticity as an adaptation to population-level environmental uncertainty used as its inspiration the life cycle of desert annual plants (Cohen, 1966). Because some years lack sufficient rain to make successful germination possible, every seed that germinates in that year is virtually guaranteed to die without reproducing. Many species seem to have adapted by delaying germination in some seeds, so that even if all germinating seeds fail to reproduce in one year, others remain dormant and can try the next year. Cohen showed that as the frequency of years unsuitable for germination increases, the fraction of seeds germinating should decrease. This relationship is supported by a number of field and experimental studies (e.g. Philippi, 1993; Clauss and Venable, 2000; Venable, 2007).

1.2 Predictive plasticity as an adaptation to environmental information

Any feature of the environment that directly causes variation in fitness contributes to environmental risk. However, other features of the environment—environmental cues—though they may not directly affect fitness, can nonetheless help individuals predict environmental change. In some cases, an individual can reduce the impact of environmental uncertainty by using a cue to alter its behavior or morphology in a way that increases its chances of survival. If this kind of plasticity is only possible during development, it will produce variation between individuals, while if it continues throughout life, it will produce variation within individuals. Which of these possibilities occurs will depend on the time lag from cue to response, and how much information is available given such a time lag (Padilla and Adolph, 1996; DeWitt et al., 1998; Gabriel et al., 2005).

Lifelong, reversible plasticity may be adaptive if environmental conditions vary within an individual's lifetime. This may happen because the environment fluctuates on a time scale shorter

than generation time, or because individuals move between different patches of habitat during their lifetimes (Levins's 1968 "fine-grained" environmental variation.) For example, several species of Darwin's finches show marked variation in both diet and foraging technique depending both on the season and the type of habitat (Tebbich et al., 2004). Each individual is a generalist, because it can respond effectively to a wide variety of circumstances. This kind of flexibility can be very powerful, but may be constrained if the organism cannot effectively predict environmental change, or if it cannot respond quickly enough.

Developmental plasticity may be adaptive if environmental conditions, though variable, are fairly stable within an individual's lifetime. This may occur if the environment fluctuates on a time scale similar to that of generations, or if individuals disperse just once and then settle (Levins's 1968 "coarse-grained" environmental variation.) Many traits, though fixed in adulthood, can be more easily repatterned during development. As long as cues received during development have some predictive value, the ability to modify the developmental trajectory can be advantageous. A classic example of adaptive developmental plasticity is shell shape in acorn barnacles; larvae settling in regions populated by a predatory snail develop a special bent morphology that is resistant to predation (Lively, 1986b).

Because behavioral traits are generally more labile than morphological traits, behavioral traits are more likely to show adaptive plasticity within individuals, while morphological traits are more likely to show adaptive plasticity between individuals (West-Eberhard, 1989; van Buskirk, 2002). However, each of these two rules also shows exceptions. Some behavioral traits are learned during a sensitive period and remain fairly inflexible afterwards. The exceptional niche breadth of the Cocos Island finch is partially due to variation between individuals in learned foraging techniques (Werner and Sherry, 1987). On the other hand, not all morphological traits are irreversible. For example, tadpoles can alter their body proportions in response to the addition or removal of predators throughout ontogeny (Relyea, 2003).

1.3 The impact of environmental variation depends on fitness trade-offs

A fitness trade-off exists between two different environments when being optimally suited to both conditions at the same time is impossible. For example, freshwater snails with rounded shells are

less easily crushed by predatory fish, while those with elongated shells are less easily pried out by crayfish (DeWitt et al., 2000). Any individual snail can thus only be well-protected against predation by one of these two major predators: a trade-off exists between defense against one and defense against the other. Which strategy is best: an extreme shell shape that is very well protected against one type, an intermediate shell shape that is somewhat protected against both, or a mixture of the two extreme shell shapes? Levins introduced the notion of the fitness set (Levins, 1962, 1968) to address this kind of question. He showed, in a graphical and intuitive fashion, how the strength of the fitness trade-offs and the structure of environmental risk interact to determine whether a specialist, a generalist, or a mixture of specialists is optimal (see Chapter 3 for an in-depth review.)

1.4 An outline of the dissertation

When individuals can react flexibly to environmental events, but not all individuals receive the same information about those events, communication may be beneficial. For example, vervet monkeys encountering three different classes of predators use qualitatively different alarm calls to distinguish between them (Seyfarth et al., 1980). In contrast, ground squirrels use different calls to indicate the urgency of a situation; one type of call is given in response to most ground predators and distant aerial predators, while the other type is given in response to most aerial predators and fast-approaching ground predators (Robinson, 1981). Why do some alarm calling systems distinguish between types of predators, while others distinguish only the level of urgency? In Chapter 2 (previously published as Donaldson et al. 2007), I use a game-theoretic model of the evolution of communication to show that asymmetric fitness trade-offs are key. Stable communication systems tend to group together situations that require similar responses. Which kind of alarm calling system evolves will depend on whether these groupings fit more naturally with different types of predators, or with the urgency of response.

The fitness set framework developed by Levins (1962, 1968) can be used to show that a combination of strong fitness trade-offs and population-level risk can make bet-hedging adaptive. How much random phenotypic diversity would the optimal bet-hedging strategy produce? For the very special case when each phenotype can only survive in one environment, the proportions of each phenotype should exactly match the proportions of the different environments (Cohen, 1966; Bergstrom

and Lachmann, 2004; Kussell and Leibler, 2005). In Chapter 3 (previously published as Donaldson-Matasci et al. 2008, footnotes added), I extend Levins's framework to show how more general fitness trade-offs determine the relationship between the amount of environmental uncertainty and adaptive phenotypic diversity. The basic idea is that each phenotype may be seen as investing a certain proportion of its reproductive effort in each of the possible environments. A bet-hedging lineage that produces the phenotypes in just the right proportions—so that the overall reproductive investment in each environment matches the environmental frequencies—grows faster on average than other lineages.

Developmental cues can provide organisms with a source of information about the future environment. This information can be used to conditionally regulate development, allowing better phenotype-environment matching. However, if the cue is not completely reliable, it is also a source of uncertainty. The effect of that uncertainty depends on how it is structured: are all individuals in a single generation subject to the same developmental cues, or do some individuals receive different cues, even though the environment they will experience is identical? In Chapter 4, I use the framework developed in Chapter 3 to analyze how the structure of error in a cue impacts the evolution of adaptive developmental strategies. I argue that the value of a developmental cue depends not simply on the amount of information it provides to a single individual about the environment, but, more importantly, on the total amount of information provided to all individuals in the population.

Chapter 2

BEHAVIORAL FLEXIBILITY AND COMMUNICATION: HOW FITNESS TRADE-OFFS SHAPE ANIMAL SIGNALLING SYSTEMS

2.1 Introduction

Most of the signals that animals use to communicate with one another do not seem to have a specific meaning in the same sense that nouns in human language do. Rather, these signals communicate about the intentions, emotional state, or identity of the sender. For example, the song of a male Darwin's finch is thought to identify him as such to conspecific females (Grant and Grant, 1997). Little blue penguins use calls to signal their readiness to escalate a fight (Waas, 1991). Even the alarm calls given by ground squirrels, which were once thought to indicate the type of predator, have been shown instead to relate to the degree of urgency perceived by the caller (Robinson, 1981). However, there are other animal communication systems in which the signals really do seem to refer to some external stimulus. Most famously, vervet monkeys use three qualitatively different alarm calls to distinguish between leopards, eagles and snakes (Cheney and Seyfarth, 1990). Similar predator-specific calls have been found in other primate species (Macedonia, 1990; Zuberbühler et al., 1997) as well as suricates (Manser, 2001) and chickens (Evans et al., 1993). This type of system is not limited to predator warnings alone: toque macaques (Dittus, 1984) and chickens (Evans and Evans, 1999) produce specific calls which alert others to the presence of food.

Semantic communication has been suggested as one of the fundamental differences between animals and humans (e.g. Bickerton, 1990). The communication systems described above, though relatively rare, are of special significance because they hint at the ability of animals to communicate about external objects and events. But does a leopard alarm call really refer to a leopard, in the same sense that the word "leopard" does? Philosophers of language contend that understanding how an utterance is used is insufficient to determine its meaning (Grice, 1957; Quine, 1960); according to this view we can never discover the true meaning of any animal signal. Ethologists have instead focused on demonstrating that some animal signals have the property of *functional reference*: the

way in which they are used, and the responses that they engender, give the appearance of referring to some external stimulus (Marler et al., 1992; Macedonia and Evans, 1993). The notion that animal signals may have some external referent is not diametrically opposed to the idea that they convey motivational information; rather, it is now well recognized that, like human language, animal signals may simultaneously do both. Still, it is possible to differentiate between systems like the vervet monkeys', which primarily refer to external objects, and systems like the ground squirrels', which primarily reflect the degree of urgency; we are interested in the evolutionary reasons behind this kind of difference.

In this report, we present a model for the evolution of functionally referential meaning in animal communication systems. We begin with a simple action-response model in which selective pressure on the production of the signal is produced by the reactions of those who respond to it, and vice versa. Selection on signals and selection on responses will often work towards one another, eventually leading to a stable and coherent communication system, as has been demonstrated previously with similar models (Hurford, 1989; Wärneryd, 1993; Nowak and Krakauer, 1999). However, these models invariably assume that the world itself takes on a very simple structure: each situation requires a particular, unique response, and all possible alternatives are equally inappropriate. Although this may be an adequate representation of certain economic games, it does not describe animal signalling interactions very well. For example, when a vervet monkey is approached by a leopard, the typical response to an eagle—looking up and running into cover—is much more dangerous than the typical response to a snake—scanning the area (Seyfarth et al., 1980).

In our model of communication in a “structured world”, we are able to represent the distinction between not-quite-optimal actions and utterly disastrous ones. We find that a wider variety of signalling systems are evolutionarily stable in our model than in the unstructured worlds of previous models, and this diversity of equilibria more accurately reflects the diversity of modern animal communication. In addition, our model suggests that evolved communication systems may facilitate the categorization of events or situations by appropriate responses, rather than by shared physical characteristics. This may explain why primarily motivational alarm call systems, like that of ground squirrels, are so common, while primarily referential ones, like the vervets', are relatively rare. If motivational states (like fear, arousal, or hunger) have evolved to help organisms make advantageous decisions, then in many cases they may be sufficient to predict an appropriate response to the

situation, and thus sufficient to determine which signal to produce. Only in special cases, where the possible reactions are too complex to be determined simply by the urgency of the situation, will a system evolve the characteristic of functional reference.

2.2 A model for the evolution of communication

Since we are interested in modeling the way that a signal, through use, may come to represent an object or a situation, we begin with a simple sender-receiver game. One individual responds to a stimulus in some observable way; another individual observes that response and reacts in turn. The first individual's action has no power to affect her payoff, while the second individual's reaction affects the payoff of both. In this sense, the first individual's action may be seen as a potential signal to the second individual; it is only through natural selection that these actions gain the status of true signals, as defined by Maynard Smith and Harper (2003, pg. 3): “an act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved.” Once natural selection begins to shape the behavior of individuals in both roles, all of the potential signals that are in use become real signals. Some of these signals may later fall out of use, preventing selection on the response. However, as long as some tendency to respond remains—however it may change through drift—they retain their power to be used as signals.

Now we can define the game more rigorously. The first player, the *signaller*, observes the state of the world $t \in \mathcal{T} = \{t_1, t_2, \dots, t_l\}$, and selects a signal $s \in \mathcal{S} = \{s_1, s_2, \dots, s_m\}$. The second player, the *signal receiver*, does not know the state of the world directly, but instead observes the signal s and chooses an action $a \in \mathcal{A} = \{a_1, a_2, \dots, a_n\}$. Note that the number of distinct signals, m , may be different from the number of states, l , or the number of possible actions, n ; we discuss the biological factors affecting the relative numbers of each at the end of this section. We will (conventionally, if somewhat unrealistically (Lachmann et al., 2001)) assume a purely cooperative game: both signaller and receiver obtain the same payoff $\pi(t, a)$, which depends only on the state of the world and the selected response. Since the payoffs are independent of the signal used, all signals are in this sense equivalent to one another. For simplicity, we also assume that all signals are transmitted without error.

In this sender-receiver game, the signaller's strategy can be represented by a matrix \mathbf{P} which contains the conditional probabilities $p(s|t)$ of producing each signal s , given each world state t . Similarly, the receiver's strategy is represented as a matrix \mathbf{Q} that provides the conditional probabilities $q(a|s)$ of selecting an action a , given signal s . Each individual can play both signalling and receiving roles, so a complete strategy R consists of both a \mathbf{P} matrix and a \mathbf{Q} matrix.

We can calculate expected payoffs, given a probability distribution on world states $p(t)$. If we further assume that each individual spends half the time as signaller and half the time as receiver, the expected payoff to an individual with strategy $R = (\mathbf{P}, \mathbf{Q})$ of interacting with an individual with strategy $R' = (\mathbf{P}', \mathbf{Q}')$ will be:

$$\begin{aligned} \bar{\pi}(R, R') &= \frac{1}{2} \sum_{t \in \mathcal{T}} \sum_{s \in \mathcal{S}} \sum_{a \in \mathcal{A}} p(t) p(s|t) q'(a|s) \pi(t, a) \\ &\quad + \frac{1}{2} \sum_{t \in \mathcal{T}} \sum_{s \in \mathcal{S}} \sum_{a \in \mathcal{A}} p(t) p'(s|t) q(a|s) \pi(t, a). \end{aligned}$$

Here the first summand is the expected payoff to the individual when acting as a signaller, and the second summand is her expected payoff as a receiver.

Since an individual's ability to communicate depends not only on her own strategy, but also on those of others around her, selection for communicative ability is frequency-dependent. Evolutionary game theory gives us a way to model these sorts of interactions. In particular, the concept of an *evolutionarily stable strategy* (ESS) (Maynard Smith and Price, 1973) provides a way to characterize the long-term behavior of a population without getting into the details of the evolutionary dynamics away from equilibrium. We assume that individuals reproduce asexually according to fitness, which is proportional to payoff in the game, and that offspring learn or otherwise inherit the strategy of their parents with some chance of error. Under a wide range of evolutionary dynamics, a population of individuals playing an ESS will be invulnerable to invasion by mutant strategies. In fact, such a population is also a long-term endpoint of the evolutionary process (Eshel and Feldman, 1984; Hammerstein, 1996).

Lewis (1969) uses a similar signaller-receiver game to describe the conventionalization of meaning in natural language, but he does not address evolutionary questions. Hurford (1989) uses computer simulations to look at the evolutionary process and Trapa and Nowak (2000) find the evolu-

Π	a_1	a_2	a_3
t_1	1	0	0
t_2	0	1	0
t_3	0	0	1

(a)

Π	a_1	a_2	a_3	a_4
t_1	8	1	1	3
t_2	0	7	1	3
t_3	0	0	6	3

(b)

Figure 2.1: Sample payoff matrices for (a) the Hurford-Nowak model (Hurford, 1989; Nowak and Krakauer, 1999; Trapa and Nowak, 2000) and (b) our model. In (a), each state has one appropriate response, and all others are useless. In (b), actions a_1 , a_2 , and a_3 are best responses to t_1 , t_2 , and t_3 but general-purpose action a_4 may be better when the state of the world is uncertain.

tionarily stable states of a related model. These models differ from ours in that the receivers choose an interpretation from the original set of world states, rather than choosing an action. Fitness is determined by the proportion of correct interpretations. However, the very idea that signals have interpretations presupposes that the communication system is used to convey referential meaning. Since we are interested in the evolution of reference, we prefer to extend a model developed by Wårneryd (1993). His paper is not really about the evolution of communication in itself; his primary goal is to show how cost-free, arbitrary signals can stabilize equilibria in a cooperative game. However, since it makes no assumptions about the “meaning” of a signal, it provides an ideal framework for exploring the evolution of motivational and referential communication.

Wårneryd’s framework is more general than the Hurford-Nowak model mentioned earlier, because the players respond to a signal by choosing an action, rather than simply inferring which situation gave rise to the signal. However, he assumes a special form of payoff matrix on signals and actions which makes his model functionally equivalent to theirs. In his model, each world state has a unique best response, and the payoff for all other actions is zero. We relax this assumption to permit arbitrary structural relationships between world states. This representation allows for much more realistic models of animal communication systems. While each state has an optimal response, we allow some of the remaining responses to be better than others. In addition, some actions may be reasonably good for several situations, without being ideal for any (see Figure 2.1.)

Wårneryd also assumes that there are at least as many signals as there are states or actions. We do not, and we will be particularly interested in cases where the number of signals is smaller than either. These cases seem most similar to real animal communication systems over a wide range

of taxa, in which the assortment of distinct signal types is surprisingly limited (Moynihan, 1970). Why should this be so? One limitation is imposed by the receivers, who must be not only able but also likely to perceive the signaler's action. That is, we can restrict our attention to the domain, or domains, in which actions cause others to react. This could be, for example, sounds within a certain range of frequencies, or the position of the tail feathers. Another limitation is that the receivers must be able to reliably discriminate different signals. The effects of a noisy environment can create a tradeoff between increasing the number of signals and being able to distinguish between them (Nowak et al., 1999). Since we are interested in the evolution of the function rather than the form of the signal, instead of explicitly modeling this process we will simply assume a fixed number of signals (but see Zuidema, 2003, for a computational approach to modeling both processes together.)

2.3 Evolutionary stability of communication systems

A Nash equilibrium strategy is one which is a best reply to itself; when such a strategy is common, though no alternate strategy can be selected for, some may drift in neutrally. In contrast, a *strict* Nash equilibrium strategy outperforms all other strategies when playing against itself, so no strategy can neutrally invade. The conditions for an ESS lie in between these two extremes: some strategies can invade neutrally, as long as the ESS is strictly superior once the invading strategy becomes common. So, in general, a strict Nash equilibrium is a special type of ESS. However, Selten (1980) showed that for role-asymmetric games (in which players are assigned different roles), *every* ESS must be a strict Nash equilibrium. In this game, therefore, the signalling strategy in an ESS must be uniquely optimal against the receiving strategy, and vice versa.

The following four conditions are necessary for the signalling system $R = (\mathbf{P}, \mathbf{Q})$ to be a strict Nash equilibrium, and therefore an evolutionarily stable strategy. The first two properties follow directly from Selten's (1980) proof.

Property 1. *The signalling strategy \mathbf{P} must be binary; that is, each state gives rise to exactly one signal.*

Property 2. *The receiving strategy \mathbf{Q} must be binary; that is, each signal results in exactly one action.*

Property 3. *The signalling strategy \mathbf{P} must be onto; that is, every signal must be used.*

Proof. Suppose that \mathbf{P} is not onto; the i th column in \mathbf{P} , corresponding to the production of signal s_i , is all zeros. Then the i th row in \mathbf{Q} , corresponding to the response to s_i , can be altered without changing the expected payoff. Thus \mathbf{Q} is not the unique best reply to \mathbf{P} , so R cannot be a strict Nash equilibrium. \square

Property 4. *The receiving strategy \mathbf{Q} must be one-to-one; that is, no two signals may give rise to the same action.*

Proof. Suppose that \mathbf{Q} maps two signals to the same action. Since \mathbf{Q} is binary, there must then be two identical rows in \mathbf{Q} , say those indicating the response to signals s_i and s_j . Then we can swap the i th and j th columns in \mathbf{P} , which are the production conditions for the two signals, without changing the expected payoff. The resulting \mathbf{P}' must differ from \mathbf{P} because, by Properties 1 and 3, no two columns are identical. Therefore \mathbf{P} is not the unique best reply to \mathbf{Q} , and R cannot be a strict Nash equilibrium. \square

These properties limit the multiplicity that is allowable in the signalling and receiving mappings. There are four possible types of multiplicity: (1) one situation leads to multiple signals, (2) one signal leads to multiple actions, (3) multiple situations lead to the same signal, and (4) multiple signals lead to the same action. The first two, as stated in Properties 1 and 2, are addressed by Selten's theorem: an ESS can have only one possible response to each circumstance. Even if some responses may sometimes perform better than others, as long as the player has no further information, the best she can do is to calculate the response which gives the highest payoff on average. The fourth multiplicity is also disallowed, as stated in Property 4: if more than one signal gave rise to the same action, then signallers could use the two interchangeably. However, as we saw above, using two signals in the same situation is never part of a stable strategy. The third type of multiplicity, on the other hand, is perfectly okay: if signallers use the same signal in multiple situations, the signal comes to "mean" to the receivers that one of several situations has occurred, each with some specified probability. As long as the payoffs are asymmetrical, it is still possible to calculate the action with maximal payoff.

While all evolutionarily stable communication systems must meet these four conditions, there are some systems which display these properties and yet are not stable. Next we add two additional properties which fill out the set of sufficient conditions for evolutionary stability.

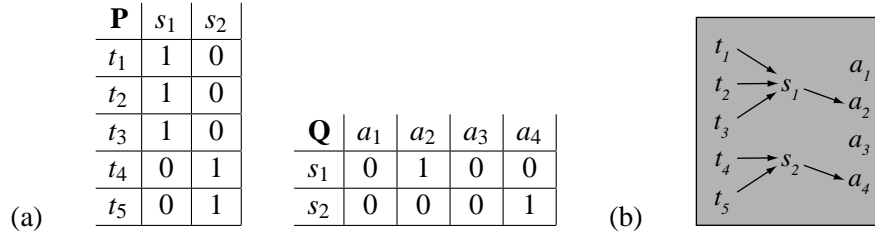


Figure 2.2: An illustration of pooling. An example strategy represented by (a) the signalling and receiving matrices, and (b) a schematic diagram. Here, signal s_1 is associated with the pool of states $\{t_1, t_2, t_3\}$ and s_2 is associated with $\{t_4, t_5\}$.

When there are more signals than states, no strategy fulfills the four conditions above; Properties 1 and 3 cannot hold simultaneously. Similarly, when there are more signals than actions, Properties 2 and 4 are in conflict. When there are equal numbers of states and signals or signals and actions, these conditions impose an exact correspondence between them.

Perhaps the most interesting case is when there are fewer signals than states, because this seems to reflect what we see in most animal communication systems today. In this case, multiple states map to a single signal, which in turn maps to just one action. This divides the set of all world states into smaller, non-overlapping *pools* (see Figure 2.2.) There is one pool for each signal, and every world state is included in some pool. Note that this usage of the term pool is similar in spirit to the notion of semi-pooling equilibria in costly signalling theory (Lachmann and Bergstrom, 1998; Bergstrom and Godfrey Smith, 1998). However, the reason for the grouping of signaller types in the costly signalling models is quite different: the conflict of interest between signaller and receiver means that in some cases a compromise can be reached in which only partial information is sent. In the current model, we assume that signallers and receivers share the same interests, so the pooling is due only to a limitation in the signals.

The following definitions and properties assume a strategy $R = (\mathbf{P}, \mathbf{Q})$ which satisfies the conditions in Properties 1-4.

Definition 1. The pool of states $\tau(s)$ associated with a signal s is the set of states mapping to that signal under \mathbf{P} : $\tau(s) = \{t : p(s|t) = 1\}$.

Definition 2. A best response to a pool of states is an action which maximizes the expected payoff

for all members of the pool:

$$\text{BR}(\tau) = \arg \max_{a \in \mathcal{A}} \sum_{t \in \tau} p(t) \pi(t, a).$$

If there is a unique such action, it is termed the *strict best response* (SBR) to the pool.

Property 5. *Every pool must have a strict best response, and the signal corresponding to that pool must map to it: $q(\text{BR}(\tau(s))|s) = 1$ for all $s \in \mathcal{S}$.*

Proof. If there is a strict best response to a pool of states grouped under one signal by \mathbf{P} , clearly an optimal reply \mathbf{Q} must map the signal to that action. On the other hand, suppose there is no strict best response to a pool $\tau(s)$. Then at least two different actions maximize the expected payoff for the pool. Any one of these actions can be chosen as the response to the signal s without changing the overall payoff. Thus there is no strict optimal reply to \mathbf{P} , and R cannot be a strict Nash equilibrium strategy. \square

Property 6. *For each member of a pool of states, the SBR for that pool must be a better response than the SBR of any other pool. That is, for all $t \in \tau(s_i)$ and $s_j \neq s_i$,*

$$\pi(t, \text{BR}(\tau(s_i))) > \pi(t, \text{BR}(\tau(s_j))).$$

Proof. We assume every signal maps to the strict best response for its pool of states. Suppose that for one state t within a pool $\tau(s_i)$, the SBR of another pool $\tau(s_j)$ provides an equally good or better response; then the signalling strategy can be changed to map t to s_j instead of s_i . The resulting \mathbf{P}' will perform just as well or better against \mathbf{Q} than \mathbf{P} does, so R cannot be a strict Nash equilibrium. \square

Theorem 1. *A strategy $R = (\mathbf{P}, \mathbf{Q})$ is evolutionarily stable if and only if the six properties listed above hold.*

Proof. We have already shown necessity; we now show sufficiency. Given that Properties 1-4 hold, Property 5 ensures that \mathbf{Q} is the single best reply strategy to \mathbf{P} , and Property 6 ensures that \mathbf{P} is the single best reply strategy to \mathbf{Q} . Therefore, any other strategy will do strictly worse against R than R does against itself, so R must be an evolutionarily stable strategy. \square

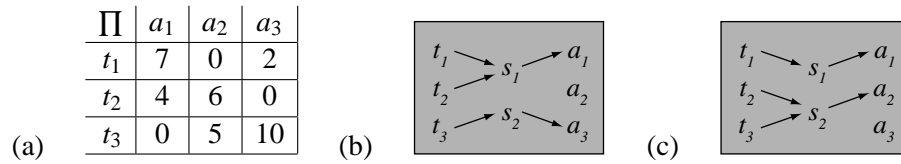


Figure 2.3: Different poolings can yield different ESSs for the same system. In this example, all states are equally likely. The evolutionarily stable communication system shown in (b) has a payoff of 7 and is payoff-maximizing, while the evolutionarily stable communication system shown in (c) has a payoff of 6. The third possible pooling, not shown, is not evolutionarily stable.

Example 1. Consider the case where there are equal numbers of states, signals and actions, and each state has a unique best response. Then an evolutionarily stable strategy will assign a signal to each state, and map that signal to the state's best response. The assignment of signal to state is arbitrary, so there will be one such ESS for every possible permutation of signals; functionally, however, all these strategies are equivalent.

In the papers by Wärneryd (1993) and Trapa and Nowak (2000), such communication systems are the only possible evolutionarily stable strategies. Constraining the probability of states and the payoff matrix to be completely symmetrical means that no pool of states bigger than one can have a unique best response; this means that under such a model an ESS cannot exist if there are fewer signals than states. In real biological systems, however, two different things are almost never exactly equally likely, nor do they give exactly the same fitness. In a model which removes these unrealistic constraints, we will see that evolutionary stability is not only possible, but even likely.

Example 2. Consider the system shown in Figure 2.3. When there are fewer signals than states, an ESS must group some of the states together. The most efficient grouping, shown in Figure 2.3b, maps the states t_1 and t_2 to a single signal, while t_3 is differentiated from the others. This is an evolutionarily stable strategy because both pooling properties are satisfied. While a_1 is not the optimal action for t_2 , it is the best of the limited possibilities created by the receiver's strategy.

The communication system illustrated in Figure 2.3c is also stable, but non-optimal. The pooling in this strategy creates a kind of evolutionary cul-de-sac: no improvement is possible by changes in either the signalling strategy or the receiving strategy alone.

In some cases, there may be no strategy which is strictly superior to all invaders. What then

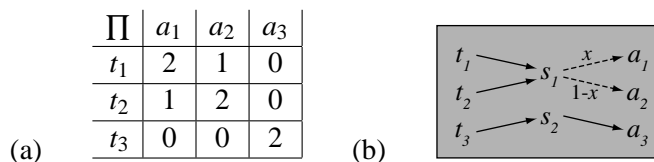


Figure 2.4: An example of a system with an ES set consisting of a single signalling strategy and multiple receiving strategies. The payoff matrix is displayed in (a), and all states are equally likely. The strategy displayed in (b) is payoff-maximizing for any $x \in [0, 1]$, where x is the probability of a response a_1 to the signal s_1 . The set of all such strategies is evolutionarily stable.

can we expect to happen, after enough evolutionary time? One possibility is that a *set* of strategies exists such that any strategy in the set is neutrally stable with regard to any other strategy in the set, but which as a set is invulnerable to invasion from outside the set. This is called an *evolutionarily stable set*, or ES set (Thomas, 1985; Balkenborg and Schlag, 2001). In this case, the system never reaches a true equilibrium, but can drift neutrally among the strategies in the set without leaving.

The conditions described above for an ESS need be modified only minimally in order to characterize an evolutionarily stable set. Rather than requiring a single uninvadable strategy, we look for a set of strategies, all of which generate the same payoff against one another, but which are otherwise uninvadable. Just as the role asymmetry in the game ensures that any ESS must be a strict Nash equilibrium, it also guarantees that any ES set must be a strict equilibrium set (Balkenborg, 1994). A strict equilibrium set is a set of Nash equilibria which is closed under best replies. For this game, this means that an ES set must consist of a pair of strategy sets, where each signalling strategy has as its set of best replies the receiving strategy set, and each receiving strategy has as its set of best replies the signalling strategy set.

Example 3. Consider a system with three states and three actions, and a payoff matrix as pictured in Figure 2.4a. If there are only two signals, one might expect an optimal strategy to group t_1 and t_2 together with one signal. However, there is then no unique best response to that signal, because either a_1 or a_2 will result in the same payoff. There can therefore be no evolutionarily stable strategy grouping the first two states; in fact, no ESS exists for this system. However, the set of strategies shown in Figure 2.4b, consisting of each pair $(\mathbf{P}, \mathbf{Q}(x))$ for all x between 0 and 1, is evolutionarily stable. The set of best responses to the pool of states $\{t_1, t_2\}$ is $\{a_1, a_2\}$ and the best response to t_3 is

a_3 , so the set of matrices $\{\mathbf{Q}(x)\}$ is the best reply set to the \mathbf{P} matrix displayed. Additionally, both t_1 and t_2 are better represented by either a_1 or a_2 than by a_3 , so \mathbf{P} is the strict best reply to any member of the set of \mathbf{Q} matrices.

Just as an ES set may have multiple receiving strategies, an ES set can also have multiple signalling strategies; this occurs when more than one signal is a best response to some state. An ES set can even have both multiple signalling strategies and multiple receiving strategies. Evolutionarily stable sets thus consist of a set of signalling strategies and a set of receiving strategies, where each member of each strategy set has as its set of best replies the entire opposing strategy set. This is still a fairly restrictive condition (particularly if one demands multiple signalling strategies and multiple receiving strategies), and there is no guarantee that an ES set will exist if an ESS does not.

Even when no ES set exists, there may be a subset of the entire strategy space which is invulnerable to invasion from outside. We can construct such a set as follows: take a single strategy and add it to the set, then add its set of best reply strategies, then take each of these strategies and add its set of best replies, and so on, stopping when all best replies are already in the set. If this set is not the entire strategy space, then the population may drift neutrally along certain paths within the set, without ever leaving the set. In this sense, it may be considered evolutionarily “stable” though not an ES set.

Example 4. Consider a system with two states, three signals, and two actions. Any binary signalling strategy \mathbf{P} which differentiates the two states will have one unused signal; since this signal is never used, any action would be an appropriate response. On the other hand, any binary receiving strategy must map more than one signal to the same action; since these two signals produce the same response, either can be substituted for the other. This means that starting with a binary signalling strategy that uses two signals, the receiving strategy can wander neutrally until the unused signal has only a single response; then the signalling strategy can wander neutrally until it switches entirely to the previously unused signal, and so on (see Figure 2.5 for a specific example.)

There are six possible binary signalling strategies which distinguish both states, represented by the upper hexagon in Figure 2.5, and there are six possible binary receiving strategies which use both actions, represented by the lower hexagon. All allowable pairs in this set (a corner in one strategy, paired with any point from the neighboring line in the other strategy) have equal, maximal

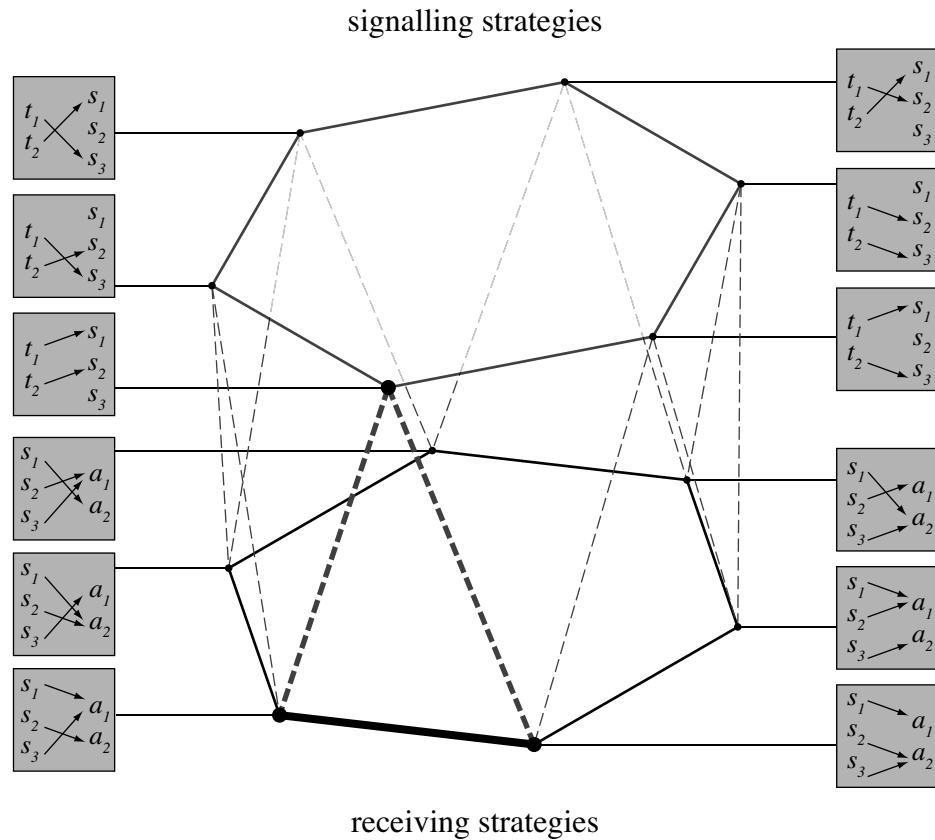


Figure 2.5: Meaning can cycle continually when there are more signals than states or actions. Here we have two states, three signals, and two actions. The upper hexagon represents a subset of signalling strategies, where each corner is a binary strategy. Adjacent binary strategies are identical except that they use different signals in response to one of the states. The line between them represents a set of strategies which map that state to the two different signals with positive probability. Similarly, the lower hexagon represents a subset of receiving strategies. When the signalling strategy is fixed at the point marked in bold, where s_3 is unused, the receiving strategy may wander neutrally along the bold line, because both actions are equally good responses to a signal that is never used. Along the interior of this line of receiving strategies, the signalling strategy remains fixed, but once either endpoint is reached (say, the one which maps s_3 to a_2) the signalling strategy will be free to vary, because now two signals (in this case, s_2 and s_3) have the same response and may therefore be used interchangeably.

fitness, so after enough evolutionary time, the system will reach some such pair. Once in this set of allowable pairs, the system will wander around neutrally, alternating between changing the signalling strategy and the receiving strategy. Notice that opposite points on the hexagons may be said to give the signals exactly opposite “meanings”; they are produced in the opposite context and induce the opposite action.

The set described in the previous example is not an ES set, because for any strategy pair within the set, most other invading strategy pairs in the set will be selected against. Only those lying along the same segment can neutrally invade. Still, once any point in the set is reached, the communication system can wander neutrally only within the set, and cannot be invaded by any strategy outside the set. This behavior is not unique to this example. We have made no mention of the payoff structure or the probability distribution of states because neither has any effect on the evolutionary behavior of the system. In fact, *any* system which has fewer states than either signals or actions will show similar behavior.

By contrast, the multiple best replies seen in Example 3 arise because two actions give equal payoffs under some pool of states. Because this type of equivalence will disappear when some of the payoffs or probabilities are changed by an arbitrarily small amount, neutral stability of this sort is unlikely to be biologically relevant. Excluding such cases, a system which is limited by the number of signals—which, we have argued, is the most biologically relevant case—will always have an evolutionarily stable strategy. Limiting the number of signals ensures that the pair of strategies which maximizes payoff will be of the form described by Properties 1–4. Without symmetry in the payoff structure, Properties 5 and 6 will hold in their strict form, so the payoff-maximizing strategy will also be an ESS.

When the number of signals exceeds the number of states and actions, every system will wander neutrally as in Example 4. In this case, a binary receiving strategy has multiple best replies because at least one signal is not used, and can therefore be responded to arbitrarily. However, if errors occur in transmission, every signal will be received with non-zero probability, and this equality in responses will no longer hold. Additionally, a binary signalling strategy has multiple best replies because two of the signals give rise to the same action, and are therefore equivalent. If the signals have different probabilities of being mistaken for one another, however, this equivalence disappears.

When error occurs and particularly when the probability of error varies from signal to signal as in Nowak et al. (1999), we could still have an ESS (that uses only some subset of the available signals) even when there are more signals than states or actions.

2.4 Conclusions

Previous models of the evolution of communication have suggested that only systems which use a unique signal for every situation can be evolutionarily stable. Under such models, if there were too few signals to distinguish all relevant cases, long-term persistence of a communication system would be impossible. Since real predator alarm call systems tend to employ only a few signals to distinguish between predators, with many types lumped together, a question remained: are these systems evolutionarily unstable, destined to change unpredictably over time, or were the models simply failing to capture some important characteristic of the system? We answer this question by extending those models to allow a more structured representation of the world: certain mistakes in comprehension may carry a higher cost than others. In doing so, we find that stable communication systems are possible under a much broader range of conditions—and thus explain how real predator alarm systems can persist over evolutionary time.

When the ability to discriminate between situations is limited by the number of signals, a communication system must group some situations together by using a common signal. In general, the stability of any particular grouping depends crucially on how important it is to distinguish among the states that are pooled together. Therefore, evolved communication systems of this sort should tend to group states which are similar in a functional sense. Rather than categorizing predators according to morphological characteristics, evolved alarm call systems should group predators which require a similar escape response. For example, Southern lapwings produce an aerial alarm call to several species of hawks, but ignore a similar-looking species which eats only fish (Walters, 1990). The notion that communication and categorization may be intimately linked by the process of evolution has been suggested before (e.g. Allen and Saidel, 1998); our paper is the first to demonstrate a mechanism for creating such a linkage.

Alarm call systems lie somewhere along a continuum between the two extremes of predator-specific systems, which distinguish between types of predators, and risk-based systems, which indi-

cate the degree of threat posed by the predator. It has been suggested that the primarily functionally referential alarm call systems of vervets and ringtailed lemurs evolved because different classes of predators require incompatible escape responses (Macedonia and Evans, 1993). Determining whether this provides a general explanation for the evolution of functional reference, of course, will require detailed study of other alarm calling systems with varying degrees of referential specificity. If the theory holds, however, our model demonstrates *why* this should be so: when categorizing situations by appropriate response yields the same groupings as categorizing them by type, a stable communication system will also show functional reference. On the other hand, when the appropriate response is dictated by the level of urgency, a stable communication system need only specify that level. Whether an alarm calling system evolves to be primarily referential or motivational is determined precisely by what types of situations require different responses.

What about systems that communicate something besides the approach of a predator—like the discovery of food, or agonistic interactions with conspecifics? Although some research has been done in both of these areas which indicates the possibility of referential communication (e.g. Hauser, 1998; Gouzoules et al., 1984) it has been more difficult to demonstrate because the responses to such calls are much less specific. In both cases, individuals react to the calls by orienting towards or approaching the caller, and what is usually measured is the latency to and/or duration of such a reaction. It has therefore been difficult to show that distinct calls really refer functionally to distinct types of food or distinct kinds of interactions. For the same reason, the hypothesis put forth for alarm calls, postulating that mutually incompatible responses to different classes of predators gives rise to referentially specific alarm calls, seems unlikely to hold here, unless there are more specific reactions to different kinds of food and/or agonistic interaction calls which we simply do not observe.

Finally, what implications does our model have for the evolution of referential communication in human language—if any? After all, since we cannot know whether animals associate their signals with some kind of internal representation of external objects, it is still possible that the kind of functional reference we have described here bears only superficial resemblance to the type of referential meaning that words in human language have (Owren and Rendall, 2001). Yet even a superficial resemblance to a referential system could have provided the conditions necessary for a truly referential system to develop. Because we make no assumptions about the “meaning” of sig-

nals, our model would provide an appropriate framework for exploring that possibility. It is still a subject of hot debate whether human language evolved from other animal communication systems for the purpose of communication, or is rather an independent outgrowth of selection for enhanced cognitive abilities (Hauser et al., 2002; Pinker and Jackendoff, 2005). Though the results described here cannot contribute directly to this debate, a model based upon ours which demonstrated how symbolic reference, as used in human language, could evolve from functional reference, as seen in other animal communication systems, would provide support at least for the plausibility of the first hypothesis.

Chapter 3

**DEVELOPMENTAL PLASTICITY AND ENVIRONMENTAL RISK:
HOW PHENOTYPIC DIVERSITY REFLECTS THE AMOUNT OF RISK****3.1 Introduction**

Organisms inhabiting a changeable, unpredictable environment face a difficult adaptive challenge. In some cases they may evolve to specialize in one particular niche, making up for poor performance in some conditions by overachieving in another. In other cases environmental uncertainty may select for a “jack-of-all-trades”, a generalist capable of overcoming any environmental hurdle but never reaching the peak productivity of a more specialized type. In this paper we focus on two different kinds of strategies that can function as adaptations to environmental uncertainty. A traditional generalist is an individual of a single, fixed type whose morphology and/or behavior is reasonably well equipped for any situation. A bet-hedging strategy, on the other hand, produces a phenotypically heterogeneous set of individuals, each of which may develop into one of several specialized types (Cooper and Kaplan, 1982; Seger and Brockmann, 1987). These two strategies may be viewed as two different ways of being a generalist: one strategy is employed by individuals, while the other is employed by genotypes. The fitness set framework introduced by Levins (1962) can be used to provide a basis for comparing the two kinds of generalists, and build intuition about the general circumstances giving rise to each.

When environmental uncertainty selects for a bet-hedging strategy, it creates non-genetic phenotypic diversity in the population. This may lead us to expect some relationship between the amount of environmental variation and the amount of phenotypic variation that is created as an adaptation to it. Such a connection has in fact been drawn for the simplest models, but a more general correspondence has been elusive (Bergstrom and Lachmann, 2004; Kussell and Leibler, 2005). We show here that the connection can be easily generalized when the phenotypes are described according to their relative fitness contributions in the different types of environments. These descriptions have a simple graphical interpretation in terms of Levins’s fitness sets, and may be interpreted as a mea-

sure of specialization. Using this measure, we develop an intuitive understanding of the quantitative relationship between environmental uncertainty and adaptive generalization.

3.2 Background: Levins's fitness sets and evolution in an uncertain environment

Whether organisms adapt to environmental uncertainty by becoming specialists or generalists—or a mixture of both—depends on a number of interacting factors (Levins, 1962; Wilson and Yoshimura, 1994). First of all, the trade-off between adaptations to different environmental conditions may take different forms. Weak trade-offs, which allow intermediate types to perform fairly well in all conditions, tend to promote the evolution of generalists. Strong trade-offs, which make intermediate types perform poorly in all conditions, tend to promote the evolution of specialists. Secondly, the structure of environmental variation plays a key role. Environmental change that affects all individuals in the population at once on the time scale of a generation, like years of drought, favors the evolution of generalists. In contrast, environmental variation between individuals in a single generation, like random dispersal into small patches of different habitats, encourages the evolution of specialists. Although some authors use the terms “temporal” and “spatial” to distinguish these two kinds of environmental variation, we prefer the more general labels *population-level* and *individual-level* environmental variation.

Levins introduced the paired concepts of the fitness set and the adaptive function as a way to gain intuition about the way that fitness trade-offs and environmental variation interact to influence the evolution of specialists and generalists (Figure 3.1). Each achievable phenotype is characterized by its fitness profile over the different environments, and then plotted on a graph whose axes are the fitness in each environment. The set of all such points in fitness space is called the *fitness set*; the portion of its boundary with negative slope is a graphical representation of the trade-off between specializing in different environments. *Weak trade-offs* create convex fitness sets, while *strong trade-offs* create concave fitness set. The *adaptive function* describes how fitness in each environment contributes to the population's overall reproductive rate. Individual-level environmental variation yields populations which reproduce according to the arithmetic mean fitness in every generation. Population-level environmental variation yields populations which vary in their productivity from generation to generation. Over the long term, such populations reproduce according to

their geometric mean fitness (Dempster, 1955; Levins, 1962; Cohen, 1966).¹ The fitness set and the adaptive function can be used together to identify a phenotype that is optimal—but only in the sense that it maximizes the population’s reproductive rate. (We will return to this point.)

One interesting result, which is easy to understand using Levins’s framework, concerns populations containing a mixture of phenotypes. The population’s reproductive rate depends on the average fitness of the mixed population in each environment. The fitness profile of the mixed population is therefore a linear combination of the fitness profiles of the phenotypes that make up the population. The set of points in fitness space that can be achieved by a mixed population is known as the *extended fitness set*. It consists of all linear combinations of the original fitness set, which considers only single phenotypes (Figure 3.1). In mathematical terminology, the extended fitness set is the *convex hull* of the fitness set. Levins distinguishes two qualitatively different cases. Where trade-offs are weak, the extended fitness set does not create any new points (Figure 3.1, right column.) Where trade-offs are strong, however, the extended fitness set does expand the boundaries. Points along this new boundary represent fitness profiles that can only be achieved with a mixture of phenotypes (Figure 3.1, left column.) When environmental uncertainty occurs at the individual level, these points will never be optimal. The contours of the adaptive function and the new boundary of the extended fitness set are both linear (Figure 3.1, upper left.) However, when environmental uncertainty occurs at the population level, the contours of the adaptive function are no longer linear. In this case, a point along the linear edge of the extended fitness set may be better than any point in the fitness set (Figure 3.1, lower left.)

Levins initially suggested that population polymorphism could therefore be an adaptation to population-level environmental uncertainty, when being a generalist is not efficient—but some care is necessary in the interpretation. If, by population polymorphism, we mean genetic polymorphism, then we must consider not only the fitness of the population compared to other populations, but also the way that natural selection within the population may act to change the mixture. Levins addressed exactly this question using a simple mendelian trait in a later paper (Levins, 1964). He also suggested that long-term selection in a fluctuating environment might act on the genetic architecture to reduce the short-term effects of selection, thus maintaining some amount of genetic variation

¹Barring complications in the genetic architecture, such as overdominance or underdominance.

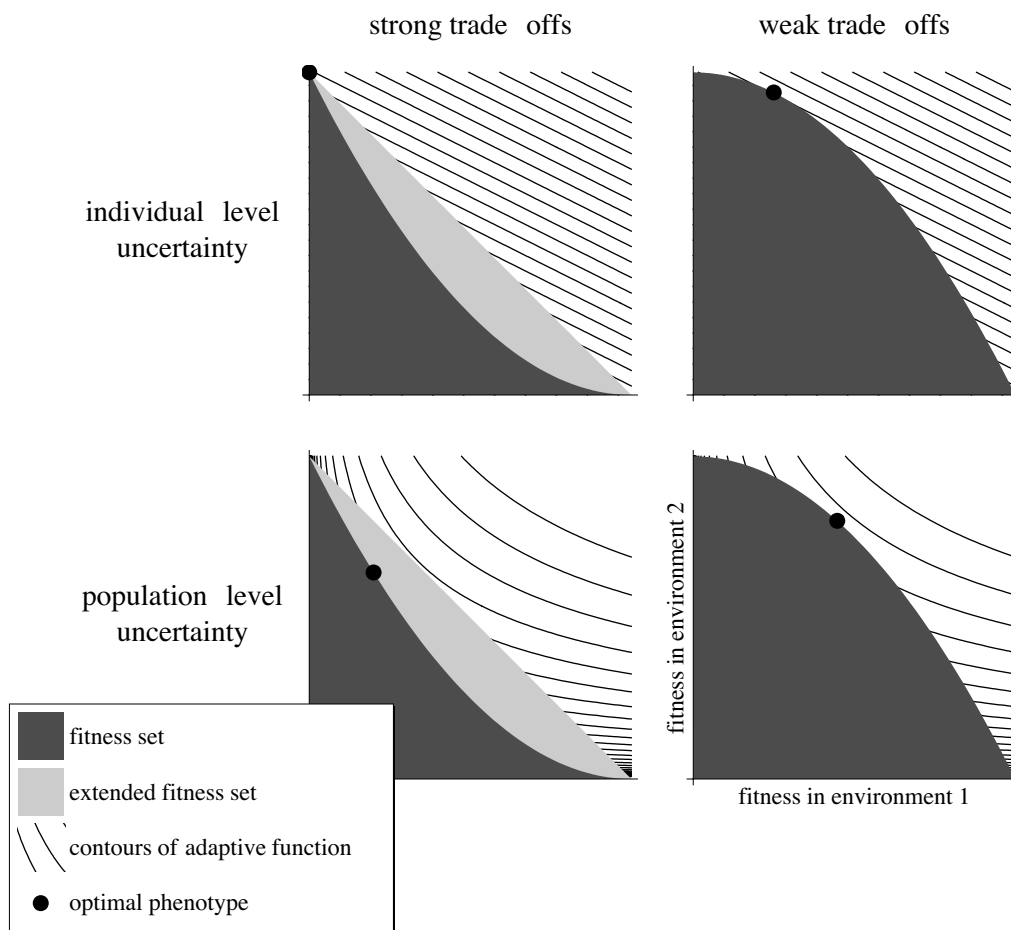


Figure 3.1: Levins's fitness set plots the fitness in each environment of all achievable phenotypes. Its boundary reflects the fitness trade-off between specializing in different environments. Trade-offs are described as weak if intermediate types can do fairly well in all environments, or strong if intermediate types do poorly in all environments. Contour lines of the adaptive function represent points in the fitness space where the population growth rates are equal. The achievable phenotype that maximizes the population growth rate is the point in the fitness set that lies on the highest contour. Specialists are favored when trade-offs are strong and environmental variation occurs at the individual level; generalists are favored when trade-offs are weak and variation occurs at the population level. The extended fitness set is created by considering in addition all mixtures of phenotypes. Allowing a mixed-phenotype population can increase the reproductive rate only when trade-offs are strong and environmental variation occurs at the population level (lower left.)

(Levins, 1965, 1968). Further work on the maintenance of genetic polymorphism as an adaptation to environmental uncertainty suggests that it is possible under certain conditions (Haldane and Jayakar, 1962; Gillespie, 1973; Sasaki and Ellner, 1995; Leimar, 2005). However, since natural selection does not generally maximize the geometric mean population fitness, Levins's fitness set framework is poorly suited to this kind of analysis (Seger and Brockmann, 1987; Godfrey-Smith, 1996).

On the other hand, phenotypic diversity in a population need not reflect genetic polymorphism. The notion of a stochastic developmental switch, which randomly produces one of several possible phenotypes, has long been recognized as a potential mechanism for producing adaptive variation (e.g. Levins, 1968), and is central to the biological theory of bet-hedging. According to this theory, organisms may adapt to population-level uncertainty in their environment by randomly developing into one of several alternative phenotypes (Cohen, 1966; Cooper and Kaplan, 1982). Such a genotype may be thought of as a “developmental generalist”, because it produces a lineage which survives well in a variety of circumstances; depending on how strong trade-offs are, it may outcompete a phenotypic generalist (Wilson and Yoshimura, 1994). Because each point within the extended fitness set may be achieved by a single genotype that produces a mixture of phenotypes, and the genotype which produces the fastest-growing lineage is likely to outcompete all others, Levins's diagrams are perfectly suited to distinguishing which of the two kinds of generalist is more efficient (Seger and Brockmann, 1987; Godfrey-Smith, 1996).

3.3 A model of the evolution of generalists

Consider a population of organisms with nonoverlapping generations. In each generation the environmental state k is drawn from some fixed probability distribution p_k , independently of the state in previous years. Each individual has a phenotype x that is fixed during development, and belongs to a continuous set of achievable phenotypes. Reproductive success depends both on the phenotype of the individual and the state of the environment. There may be individual-level variation in reproductive success within generations, but the average reproductive success of a phenotype, f_{xk} , must be consistent between generations with the same environmental state.

This is the classical model of evolution in fluctuating environments, where the phenotype with

the highest geometric mean fitness is most likely to become fixed (e.g. Dempster, 1955; Levins, 1962; Cohen, 1966). One way to understand why natural selection tends to maximize the geometric mean in such cases is to look at long sequences of environments, and ask which genotype will take over the population in most of these sequences. Under the assumptions of our model, the genotype that will win in any particular sequence of environments is the one that had the largest growth rate. For a simple fluctuating environment, the strong law of large numbers implies that in almost all long sequences, each environmental condition is experienced approximately in proportion to its probability of occurring. In such sequences of environments, the expected long-term growth rate for a lineage expressing phenotype x is the growth rate averaged over the environmental probabilities:

$$r(x) = \sum_k p_k \log f_{xk}, \quad (3.1)$$

which is the log of the geometric mean fitness. A natural extension to this approach also considers bet-hedging genotypes, which produce offspring with phenotypes given by some probability distribution g_x . Once a lineage is common enough, its average reproductive success is simply the weighted average reproductive success of the phenotypes it produces, so the long-term growth rate is given by:

$$r(g) = \sum_k p_k \log \sum_x g_x f_{xk} \quad (3.2)$$

(e.g. Seger and Brockmann, 1987; Yoshimura and Clark, 1991). The genotype that will be observed most of the time is the one that maximizes r —since its growth rate, over almost all sequences of environments, is larger than that of any other strategy.

In general, natural selection need not lead to optimization of the long-term growth rate. The model includes a number of important simplifying assumptions. In Section 3.4, we discuss the realism of these assumptions and the consequences of relaxing them.

3.3.1 *The region of strong trade-offs*

Given a model of the fitness trade-offs, we can use Levins's fitness set diagram to graphically solve for the genotype with the highest long-term growth rate. It is easy to see that when trade-offs are

weak, the best response to population-level variation will always be a single generalist phenotype. However, when trade-offs are strong, the situation is a little more complicated. A concave fitness set may show strong trade-offs only along part of its boundary. In this case, bet-hedging will be an optimal strategy only when the optimal single phenotype lies in a region of strong trade-offs. To illustrate, we begin with a simple example inspired by the ecology of amphibian metamorphosis.

Example. Frogs laying eggs in temporary ponds face a trade-off imposed by their time to metamorphosis. Late-metamorphosing tadpoles become larger frogs with higher fitness, but confront a higher risk that the pond will dry up before metamorphosis is complete. In our simple model, tadpoles are capable of metamorphosis beginning at some threshold number of days after hatching. After this, they grow linearly with the time to metamorphosis. However, if the pond dries up before metamorphosis, the tadpole dies. Each frog lays its eggs in a different pond. In dry years, the time to drying of individual ponds is independent and normally distributed.² In wet years, the time to drying of individual ponds is also independent and normally distributed, but with a different mean and standard deviation. From this we can calculate the expected fitness of a tadpole metamorphosing after a certain number of days, in a wet or dry year (Figure 3.2, see Appendix A for details of the model.)

The region of strong trade-offs is the part of the fitness set that makes it concave (Figure 3.3.) Since the adaptive function increases monotonically with fitness in each environment, the best pure strategy must lie on the boundary of the fitness set. Similarly, the best bet-hedging strategy must lie on the boundary of the extended fitness set. Whether a single generalist phenotype or a bet-hedging genotype is optimal depends on the overlap of these boundaries. We call the part where they overlap the region of weak trade-offs. The part where they do not overlap is the region of strong trade-offs; in this area, bet-hedging can improve upon any single phenotype. In particular, if the optimal strategy is a bet-hedging one, it will consist of a mixture of the phenotypes at the endpoints of the region of strong trade-offs.

We would like to determine, first of all, when bet-hedging is optimal, and secondly, if it is optimal, what is the best mixture of phenotypes. These questions can be answered for any specific

²Strictly speaking, a normal distribution of drying times is impossible, since some ponds would dry up after a negative number of days! However, for the purpose of illustration, this is a reasonable approximation.

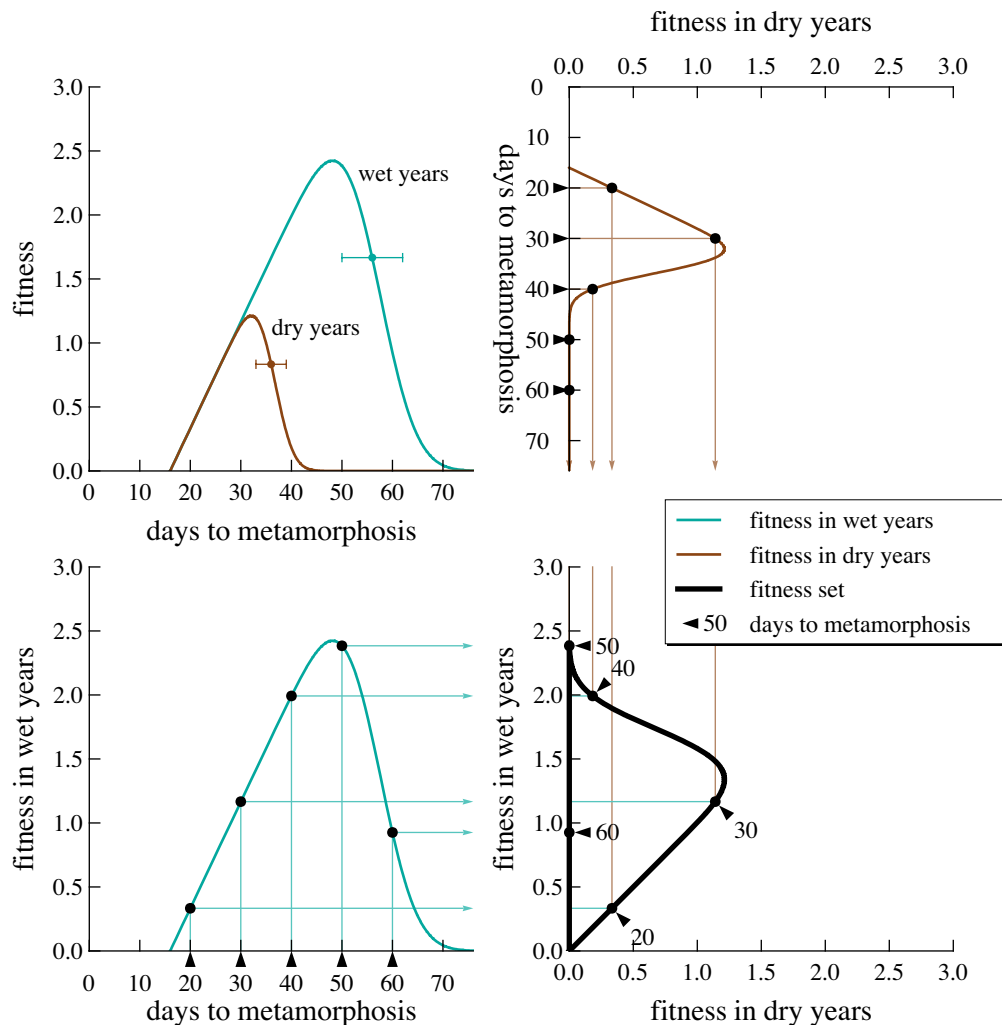


Figure 3.2: Levins's fitness set, lower right, is a parametric plot of the fitness of each phenotype in several environments. Shown on the upper left is a graph of average fitness in two kinds of years, for tadpoles that metamorphose at different ages. Error bars indicate the mean and standard deviation of time to drying. To illustrate how the parametric plot is derived, we show it alongside individual plots of fitness in each kind of year. Consider the phenotype of metamorphosis at 30 days. Its fitness in wet years, shown in the lower left plot, becomes the y coordinate in the fitness set. Its fitness in dry years, shown in the upper right plot, becomes the x coordinate in the fitness set. When this is done for every phenotype, a curve is traced out in fitness space. This is called the fitness set.

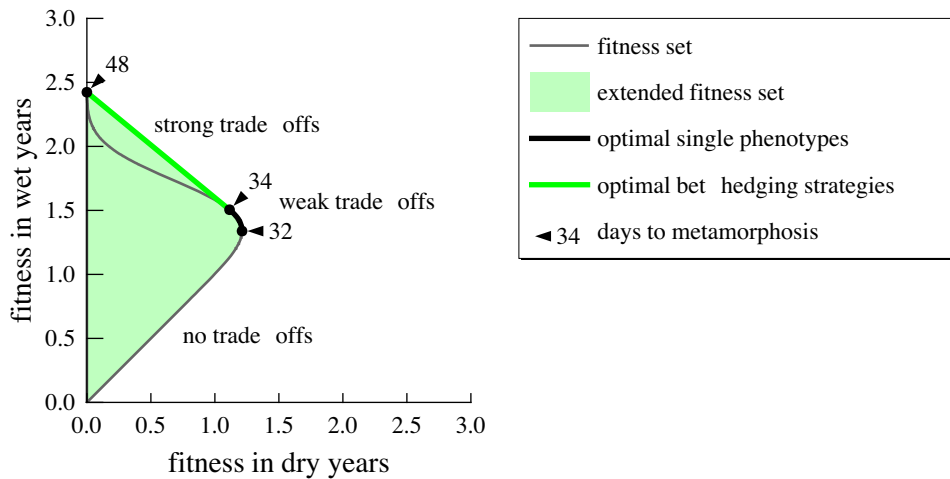


Figure 3.3: Levins's extended fitness set consists of all points in the fitness set, plus all linear combinations of those points. The boundaries of this set reflect the trade-offs in the system. Until metamorphosis at 32 days, fitness increases in both kinds of years, so there is no trade-off at all. Between 32 and 48 days, fitness in dry years decreases while fitness in wet years increases. This region of trade-offs can be divided into two parts. Between 34 and 48 days, the boundary of the extended fitness set goes beyond the boundary of the fitness set. That means that if the best pure strategy lies somewhere in this range, a mixed strategy which combines tadpoles metamorphosing at 34 and 48 days would do even better. This is called the region of strong trade-offs. Between 32 and 34 days, the boundary of the fitness set and the extended fitness set overlap. If the best pure strategy lies somewhere in this range—the region of weak trade-offs—no mixed strategy can be an improvement.

model using the traditional fitness set approach, by identifying the point of intersection between the extended fitness set and the highest contour of the adaptive function (Levins, 1968; Yoshimura and Jansen, 1996). However, a much more general and intuitive approach is possible using the measure introduced in the next section.

3.3.2 *A quantitative measure of specialization in different environments*

When each phenotype can survive in only one environment, bet-hedging is the only way to survive: a lineage of any single phenotype is sure to go extinct. The optimal bet-hedging mixture of such phenotypes—regardless of fitness differences in different environments—is to match the probability of environments. For example, Cohen’s (1966) simplest model of seed germination, in which a germinating seed can only survive in good years, showed that the optimal germination fraction is equal to the probability of a good year. The generality of this result is well known in information theory (Kelly, 1956; Cover and Thomas, 1991), but has only more recently been applied in a biological context (Bergstrom and Lachmann, 2004; Kussell and Leibler, 2005). This case provides an important reference point, and a good basis for our measure of specialization, precisely because it is so well understood. We will show that our measure can then easily be extended to treat the more general case, where phenotypes can survive in several environments.

We define a phenotype that survives in only one environment as *completely specialized* in that environment. Furthermore, a bet-hedging genotype that produces a mixture of completely specialized phenotypes is specialized in each environment according to that mixture. Consider a genotype g that allocates a fraction g_k of its offspring to a phenotype that is completely specialized in environment k . Then the fitness of that genotype in environment k is $f_{gk} = g_k d_k$, where d_k is the fitness of the completely specialized type in its environment. We say that the *degree of specialization* of the genotype in environment k is g_k , because that is the proportion of its reproductive effort that is invested in environment k . Note that, according to the result cited above, the optimal genotype should specialize in each environment according to its frequency.

The idea that a bet-hedging genotype divides its reproductive effort among environments can be extended to include phenotypes that survive in multiple environments. To illustrate how, we return to the example of amphibian metamorphosis. In the last section, we saw that any optimal bet-hedging

strategy must be a mixture of just two types: a slow-metamorphosing tadpole (at 48 days) and a fast-metamorphosing tadpole (at 34 days.) Such genotypes lie along a straight line on the boundary of the extended fitness set (Figure 3.3.) Notice, however, that this line is a subsection of a longer line that extends all the way to both axes (Figure 3.4). This longer line represents all bet-hedging strategies that combine two completely specialized types: the slow-metamorphosing type, which has fitness 2.42 in wet years, and fitness 0 in dry years, and another type that has fitness 2.94 in dry years, and fitness 0 in wet years. Since any particular tadpole would do better in a wet year than a dry one, it is impossible to be completely specialized in dry years—the second phenotype cannot actually be achieved. However, we can still ask, if it were achievable, what would be the optimal mixture of these two completely specialized types? We know the answer must be to produce dry-year specialists as often as there are dry years, and wet-year specialists as often as there are wet years. If that point lies along the subsection of the line that can be produced by combining slow and fast metamorphosing tadpoles, then the optimal allocation of reproductive effort—matching the environmental frequencies—can still be achieved.

We therefore define the degree of specialization of a bet-hedging genotype in each environment in terms of the mixture of completely specialized types to which it corresponds. Consider a bet-hedging genotype g which combines several phenotypes that are not completely specialized. Suppose that each of those phenotypes has a fitness profile that can be represented as a different mixture of just one set of completely specialized phenotypes:

$$f_{xk} = s_{xk}d_k. \quad (3.3)$$

Any bet-hedging mixture of the original phenotypes is then also equivalent to a mixture of that same set of completely specialized phenotypes:

$$\sum_x g_x f_{xk} = s_{gk}d_k, \text{ where} \quad (3.4)$$

$$s_{gk} = \sum_x g_x s_{xk} \quad (3.5)$$

We define the degree of specialization of the bet-hedging genotype in terms of this mixture, s_{gk} . Haccou and Iwasa (1995) note that, if each fitness can be written as a product $s_{xk}d_k$ as in Equation

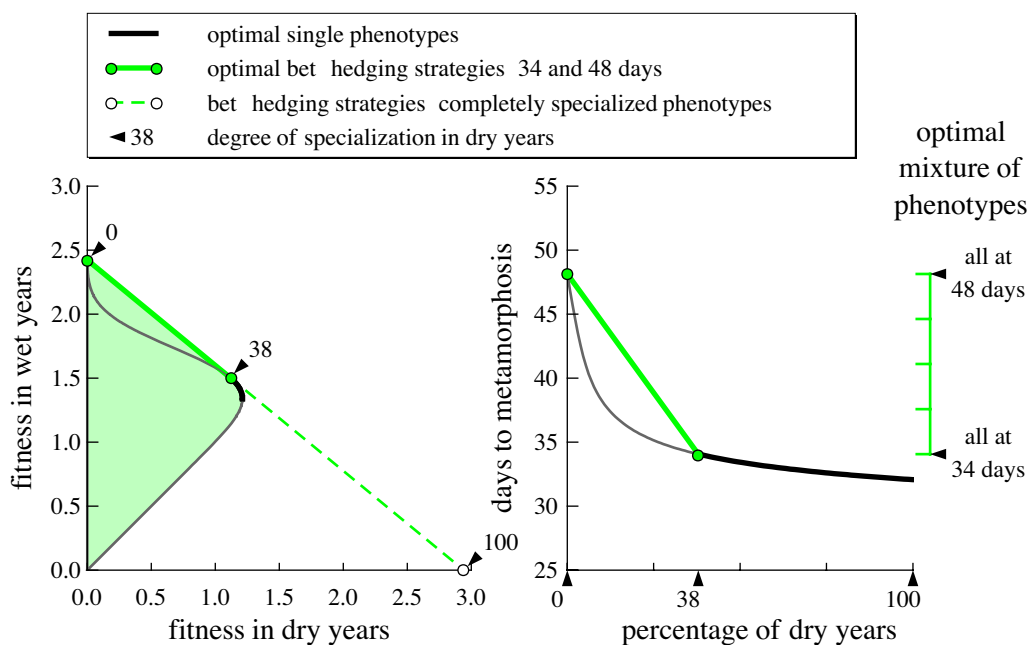


Figure 3.4: An optimal bet-hedging strategy matches its level of specialization in each environment to the frequency of that environment. In the left panel, we show how specialization levels for any bet-hedging mixture of two phenotypes can be calculated. The line between the two phenotypes plotted in fitness space is extrapolated to the axes. This defines a bet-hedging strategy that combines two perfectly specialized phenotypes. The best mixture of perfectly specialized phenotypes matches the frequency of different environments. However, in the example shown it is impossible to be perfectly specialized in dry years. Still, some of the same results can be achieved by using a mixture of the original phenotypes. In the right panel, we plot the best bet-hedging mixture of metamorphosis at 48 days and at 34 days, as a function of the percentage of dry years. The slope of this line reflects the sensitivity of the optimal strategy to changes in the environmental probabilities. The right-hand axis indicates what mixture of the two types attains the optimal level of specialization, while the left-hand axis indicates the average number of days to metamorphosis. We also show the best single-phenotype strategy, in days to metamorphosis. From 0–38% dry years, bet-hedging is better than the best single-phenotype strategy. Above this, bet-hedging cannot achieve the optimal level of specialization, so the optimal strategy is to produce a single phenotype.

3.3, then the optimal distribution of types g_x^* is the one that achieves

$$\sum_k g_x^* s_{xk} = p_k \quad (3.6)$$

where p_k is the probability distribution over environments. Our definition of specialization in terms of an equivalent mixture of completely specialized phenotypes is therefore particularly convenient, because it means the right amount of specialization in any environment, s_{g^*k} , is simply the probability of that environment.

It remains to be shown when and how the appropriate decomposition of the fitness, as in Equation 3.3, can be accomplished. When there are just two or three different kinds of environments, this is easy to visualize using Levins's fitness sets, as illustrated in Figure 3.4. We use the region of strong trade-offs to identify the phenotypes that could be used in an optimal bet-hedging strategy. The number of phenotypes in that set is limited by the number of distinct environments, and will equal the number of environments as long as there are strong fitness trade-offs between all environments. In that case, the specialization levels of the individual phenotypes s_{xk} are uniquely defined, and can be found according to the method described in Appendix B.

Looking at genotypes in terms of their specialization in different environments provides the tools we need to answer the questions posed at the end of the last section: when is bet-hedging adaptive, and if it is adaptive, how much is optimal? As shown in Figure 3.4 (left panel), a tadpole that metamorphoses at 48 days is completely specialized in wet years, while a tadpole that metamorphoses at 32 days is only 38% specialized in dry years. Therefore, if dry years occur more than 38% of the time, a bet-hedging strategy using only these two types could not be specialized enough in dry years. If dry years occur less often than this, then the optimal strategy is the combination of the two phenotypes that matches the degree of specialization in dry years to the probability of dry years (Figure 3.4, right panel). As dry years increase in probability from 0% to 38%, the optimal mixture of types changes from all metamorphosing at 48 days, to all metamorphosing at 34 days. That means that a fairly small increase in the percentage of dry years corresponds to a larger increase in the fast-metamorphosing type.

In general, bet-hedging with phenotypes that are not completely specialized in any environment can produce only a limited subset of possible specialization levels, defined by the specialization

levels of the phenotypes. When environmental probabilities lie outside this range, a single, generalist phenotype will be optimal. Within this range, on the other hand, the optimal amount of phenotypic variation changes linearly, taking on all possible values, as environmental probabilities change (see Equation 3.6). The slope of this linear relationship, reflecting the sensitivity of the optimal bet-hedging mixture to the environmental probabilities, depends only on the size of the region of strong trade-offs. The general form of the relationship between the amount of adaptive diversification and the amount of environmental uncertainty that drives it is therefore quite simple and intuitive.

An additional advantage of describing genotypes in terms of their specialization in different environments is that it allows for a simple description of the strength of selection for a bet-hedging genotype, when it is optimal. We examine the difference between the optimal growth rate, which uses the bet-hedging genotype g^* , and the current long-term growth rate under the bet-hedging genotype g :

$$r(g^*) - r(g) = \sum_k p_k \log \frac{\sum_x g_x^* f_{xk}}{\sum_x g_x f_{xk}}. \quad (3.7)$$

Rewriting the fitnesses as in Equation 3.3, and substituting in Equations 3.4 and 3.6, we find

$$r(g^*) - r(g) = \sum_k p_k \log \frac{p_k}{s_{gk}} \quad (3.8)$$

which is the Kullback-Leibler divergence between the environmental probabilities p_k and the genotypic specialization level s_{gk} —the environmental probabilities for which the current genotype g would be optimally adapted (compare Haccou and Iwasa, 1995; Kussell and Leibler, 2005). In this sense, the amount of environmental uncertainty indicates how much of a generalist an optimally adapted genotype should be; the farther away a genotype is from this ideal, the stronger the selection to improve. Furthermore, selection for moving from the best single-phenotype strategy to any bet-hedging strategy cannot exceed the Kullback-Leibler divergence between the specialization levels of the component phenotypes. If there are strong trade-offs only in a small portion of the fitness set, there will be only a narrow range of environmental uncertainty in which bet-hedging is optimal. Furthermore, selection for adopting a bet-hedging strategy will be strongly limited, because the best phenotypic generalist will do almost as well as the optimal bet-hedging genotype.

3.3.3 *The effect of individual-level risk*

Until now, we have addressed only the selective effects of population-level risk, because this is the only kind that can lead to adaptive phenotypic diversification. Now we ask, how does individual-level risk affect the strength of selection for bet-hedging that may be imposed by population-level risk? The method of dividing the variance in average genotypic fitness over generations into the variance in individual fitness and the correlation in fitness between individuals, introduced by Frank and Slatkin (1990), provides some insight into this question. They represent increasing individual-level uncertainty by decreasing correlations in fitness between individuals of the same genotype. This decreases the variance in average genotypic fitness, effectively decreasing the amount of population-level uncertainty. We may therefore expect that adding individual-level risk to our model will dampen the effects of population-level risk.

Our example of amphibian metamorphosis contains environmental risk at two different levels: at the population level, because dry years cause the ponds to dry up earlier on average, and at the individual level, because within each year the time to drying of individual ponds varies according to a normal distribution. In Levins's original formulation such combinations of risk at two different levels were represented by adjusting the adaptive function (Levins, 1962). However, we have instead included the individual-level risk in the fitness set: each phenotype is represented by a vector consisting of its average fitness in each environment. This is possible because, given the type of year, the ponds dry up independently of one another. Therefore, the variance over different years of the average genotypic fitness (a sample mean) will be negligible for large populations. In the example, individual-level uncertainty in the time to drying makes specialized types more generalist, by decreasing average fitness in the most favored circumstances and increasing it in the least favored ones. In the process, it reduces the size of the region of strong trade-offs, putting stricter and stricter limits on the amount of population-level uncertainty necessary to induce adaptive phenotypic variation. When we increase the variation within years enough that there is significant overlap in drying time between wet and dry years, the region of strong trade-offs disappears completely—making phenotypic variation non-adaptive (Figure 3.5.)

Our approach has the advantage of singling out population-level uncertainty as the driving force in the evolution of adaptive variation: it sets a target for the amount of specialization in different

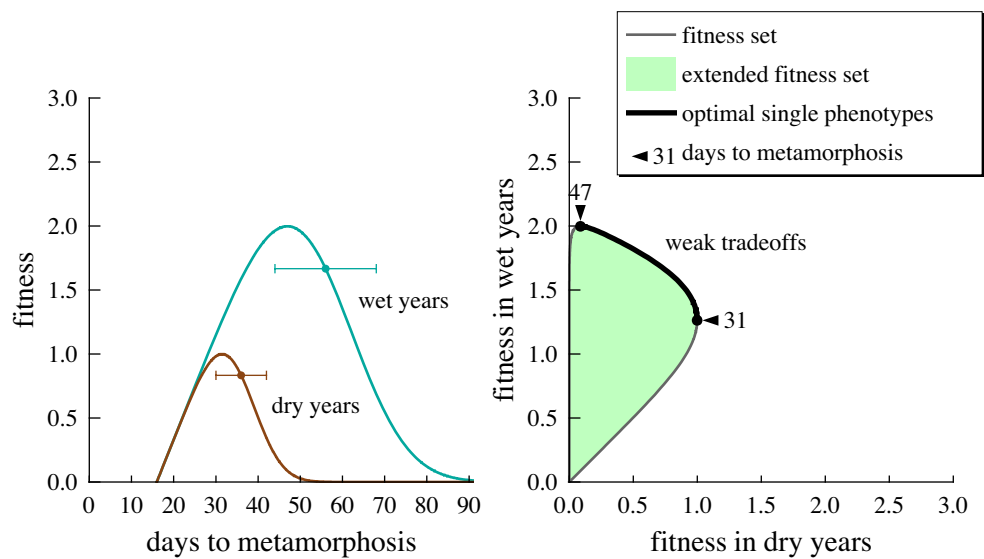


Figure 3.5: Individual-level uncertainty reduces selection for phenotypic diversification by changing the shape of the fitness set. The left panel shows fitness as a function of phenotype; the right panel shows the resulting fitness set and extended fitness set. Compared to the parameters used in Figures 3.2 and 3.3, the only difference here is a larger standard deviation in time to drying, in both kinds of years. This reduces the risk associated with being caught in the wrong kind of year, creating only a weak trade-off between specializing in wet and dry years over the entire range. In this situation, no matter the level of population-level risk, an individual-level generalist will always be favored.

environments. Individual-level uncertainty may then alter the circumstances under which diversification is favored, by changing the shape of the fitness set. Individual-level environmental risk could affect the fitness profile of different phenotypes in many ways. Its impact depends on how it affects the average performance of different types. In our example, within-year variation in time to drying of different ponds makes the average fitness of similar types more similar (the slopes in the left panel of Figure 3.5 are less steep than in the corresponding panel in Figure 3.2.) This effectively reduces the size of the region of strong trade-offs, making intermediate types better generalists, and making bet-hedging less important. On the other hand, if individual-level stochasticity affected only the variance in fitness of individual types, not the average fitness, it would not change the fitness set at all. Still, to the extent that individual-level uncertainty does make phenotypes more generalist, it will make bet-hedging both less likely to be adaptive, and less strongly selected for (see Section 3.3.2).

3.3.4 *The effect of developmental plasticity costs*

Our representation of the growth rate of a lineage stemming from a bet-hedging genotype (see Equation 3.2) implicitly assumes that there is no cost to plasticity. This assumption is important for our results, because it means that the fitness profile of a bet-hedging genotype is a simple linear combination of the fitness profiles of the phenotypes it produces. However, some kinds of cost can still be represented within our framework. If there is a fixed cost to being able to produce more than one phenotype, regardless of the proportions actually produced, the overall shape of the extended fitness set will not change. Consider a cost of plasticity that is proportional to an individual's realized fitness. This simply scales the extended fitness set by some factor (Figure 3.6, left panel.)³ The boundary of the extended fitness set will still be linear wherever there are strong trade-offs, and will still correspond to the same mixed strategies. What changes is the comparison between the best pure strategies and the best mixed strategies. Mixed strategies are optimal over a smaller range of environmental uncertainties, and the range of optimal mixtures decreases as well. The linear relationship between the environmental frequencies and the optimal mixture of phenotypes still holds (Figure 3.6, right panel.)

³However, the original fitness set does not change.

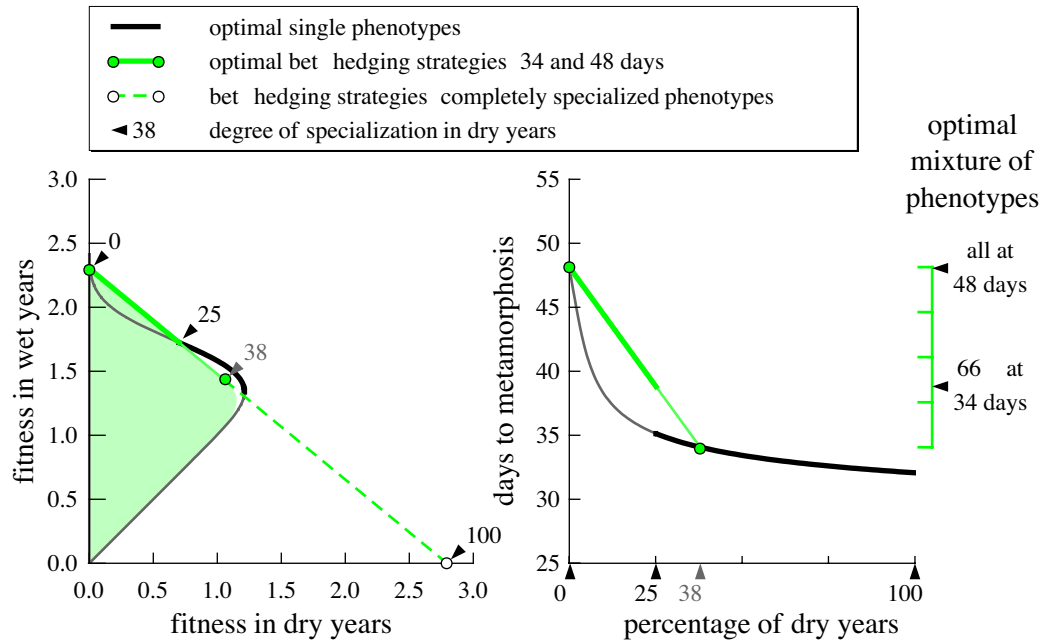


Figure 3.6: A fixed cost to plasticity reduces the range of environmental uncertainty in which bet-hedging is optimal. As in Figure 3.4, we use the extended fitness set to calculate the degree of specialization in dry years for a range of bet-hedging strategies. Because the cost is proportional to fitness, the shape of the extended fitness set does not change; it is merely scaled down. The phenotypes that are used to make all optimal bet-hedging mixtures are again metamorphosis at 34 and 48 days, and the specialization levels of each do not change. This means that the best bet-hedging strategy for any particular percentage of dry years, shown at right, also does not change. Likewise, the best single-phenotype strategy remains the same. What does change, however, is the comparison between single-phenotype strategies and bet-hedging strategies. The cost of bet-hedging means that the range of environmental uncertainty in which the best bet-hedging strategy outperforms the best single phenotype gets smaller: 0–25% dry years instead of 0–38%. This corresponds to mixtures where less than 66% of tadpoles metamorphose at 34 days. Mixtures outside this range are not worth the cost of plasticity.

Another possibility is that the cost of plasticity could depend on the amount of phenotypic diversity it generates. For example, a plastic developmental pathway might tend to build the less commonly produced phenotype with less accuracy, and thus lower average fitness. Then a bet-hedging strategy that almost always produces just one phenotype would incur very little fitness cost, while one that produces two phenotypes in equal proportions would incur a higher cost. In this case the boundary of the extended fitness set will no longer be linear. The graphical method of finding the point on the boundary which maximizes the adaptive function can still be used. However, the most important feature of our framework—the direct relationship between environmental probabilities and phenotypic diversity—is lost.

What does this mean for the applicability of our framework in natural systems? The answer depends on how common costs to plasticity are, and how those costs are related to the amount of phenotypic diversity that is created. To our knowledge, no one has yet empirically measured cost in a system where stochastic developmental plasticity has been demonstrated. Most empirical studies of the costs of plasticity have focused instead on adaptive plasticity in response to predictive cues (see DeWitt et al., 1998, for a review). Much of the cost of plasticity in these cases may be related to the ability to detect and respond appropriately to predictive cues. Such costs do not apply to stochastic developmental plasticity. However, it may still be more difficult for an organism to produce, along alternative developmental pathways, the same phenotypes that developmentally canalized organisms could produce. This difficulty may be due to limits on the range of plastic development or increased developmental instability. Empirical tests for these kinds of limitations to plasticity so far show no support for any association between increased plasticity and decreased phenotypic range or precision (e.g. DeWitt, 1998; van Kleunen et al., 2000; Relyea, 2002). This suggests that the costs to stochastic developmental plasticity may be fairly limited, and not related to the amount of stochasticity. These are precisely the conditions in which our model is most useful.

3.4 *Generality of the model*

A more formal analysis of the effects of stochastic environments on natural selection would include explicit models of population dynamics. Our assumption that the genotype with the highest average log fitness is most likely to fix in the population depends on a number of key simplifications in the

model. We consider each of these simplifications in turn.

First of all, we have made the assumption that the long-term growth rate of a lineage is a good predictor for the outcome of a competition between two genotypes. This is clearly true if both lineages are growing exponentially and completely independent of one another, but such a scenario is not biologically plausible. In most cases there will be some kind of density-dependent regulation of growth. We would like to know which strategy is likely to outcompete all others, in the long term. This depends on the way that the relative proportions of different strategies in the population change over time. Consider some form of density-dependent population regulation that is equally likely to affect all individuals, regardless of strategy or phenotype. This would maintain realistic population sizes, but would not affect the relative proportions of different strategies. Therefore, the strategy with the highest long-term growth rate would still be most likely to outcompete the others (McNamara, 1995; Grafen, 1999).

On the other hand, some kinds of density-dependent population regulation can affect the proportions of different types of strategies. For example, individuals might specialize in the use of different limiting resources, whose availability varies from generation to generation. Then an individual's reproductive fitness will depend not only on its own phenotype, the environment, and the population size, but also on the frequency of other phenotypes in the population. This is a classic case of frequency-dependent selection, whose outcome cannot generally be predicted by principles of optimization.⁴ The concept of an evolutionarily stable strategy—a genotype that, once common, resists invasion from potential mutant types—can then be a useful tool for identifying adaptive strategies (Maynard Smith, 1982). In the context of fluctuating environments, the ability of a mutant genotype to successfully invade a resident is predicted by the long-term growth rate of the mutant in a population of the resident (Metz et al., 1992; Yoshimura and Jansen, 1996). Our framework can still be used to predict what would be the best bet-hedging strategy, for a particular distribution of phenotypes in the resident population. If this optimal strategy also produces the given distribution of phenotypes, then it is at least an equilibrium strategy, although it need not be stable.

The second important assumption is our use of the log fitness, averaged over different environmental states, as a proxy for the long-term growth rate of a lineage (see Equation 3.1). Whether this

⁴In fact, fitness depends not just on frequencies, but also on total densities. Only in the classical case of “soft selection”, as considered by Levene (1953), is the fitness independent of densities.

is a good approximation depends crucially on how the environment varies on several scales: within individuals, between individuals in the same generation, and between generations.

Our assumption that individuals live and reproduce in only one generation, and that each generation is characterized by a single environmental state, ignores the possibility of environmental variation within an individual's lifetime. In general, the effect of such within-individual variation is to reduce the impact of between-generation variation, because individual lifetime reproductive success becomes less variable (Frank and Slatkin, 1990; Sasaki and Ellner, 1995). In fact, organisms may in some cases evolve to survive and reproduce over multiple seasons as a response to an environment that fluctuates from generation to generation. Evolution of such a life history strategy may therefore sometimes be seen as a kind of "risk spreading" in its own right (Murphy, 1968; Goodman, 1984).

Variation between individuals in the same generation need not affect the long-term growth rate, as long as the average fitness of each phenotype within each environmental state does not vary much. This will be true as long as the number of individuals of each phenotype is fairly large, and the reproductive success of different individuals is independent, conditional on the environmental state (Frank and Slatkin, 1990). We have included individual-level variation of exactly this sort in our example (see Section 3.3.3).

In this study, we have focused on the evolutionary impact of environmental variation between generations. We have made the important simplification that the environmental conditions in each generation do not depend on previous environmental history. In such cases, the asymptotic growth rate of a lineage simplifies to its average log fitness over generations. However, this simplification continues to hold even in more complicated scenarios: as long as the population is unstructured, and the environmental states constitute an ergodic process, the asymptotic growth rate is the log fitness of the lineage, averaged over the stationary distribution of environmental states (Tuljapurkar, 1990).

3.5 Discussion

Variable life history strategies have been studied as a potential adaptation to fluctuating environments in a wide variety of biological systems. The best-known example is delayed germination

in desert annual plants. This was the inspiration for Cohen's (1966) model, and has spawned a series of follow-ups (e.g. Philippi, 1993; Clauss and Venable, 2000; Evans et al., 2007; Venable, 2007). A similar phenomenon—an overwinter diapause—is observed in many insects and crustaceans (Saiah and Perrin, 1990; Bradford and Roff, 1997; Danforth, 1999; Hopper, 1999; Menu et al., 2000; Philippi et al., 2001). A highly variable time to metamorphosis, as described in our example, has been observed in some anurans breeding in temporary pools (Lane and Mahony, 2002; Morey and Reznick, 2004). In fish and amphibians, the trade-off between egg size and egg number may make variation in egg size an adaptation to environmental uncertainty (Capinera, 1979; Crump, 1981; Kaplan and Cooper, 1984; Koops et al., 2003). For organisms that switch between sexual and asexual modes of reproduction, like aphids and some plants, the timing of that switch may vary in response to uncertainty (Berg and Redbo-Torstensson, 1998; Halkett et al., 2004). Bacteria are rapidly becoming an important model system for the study of adaptive non-genetic phenotypic diversity, in part because the regulatory mechanisms underlying the heterogeneity are particularly amenable to study (reviewed in Avery, 2006; Smits et al., 2006). For example, *Escherichia coli* periodically exposed to antibiotics switch stochastically between a fast-growing antibiotic-sensitive type and a slow-growing, antibiotic-resistant type; the rate of switching varies between strains and may be adapted for different frequencies of antibiotics exposure (Kussell and Leibler, 2005). Finally, in a few systems it is argued that a single, low-risk life history strategy is a better adaptation to fluctuating environments than phenotypic diversity would be (Boyce and Perrins, 1987; Einum and Fleming, 2004; Hassall et al., 2006; Simons and Johnston, 2003). This is sometimes called conservative bet-hedging, in contrast to diversified bet-hedging, which uses a variety of phenotypes (Seger and Brockmann, 1987; Philippi and Seger, 1989).

Empirical studies of life history evolution in response to environmental uncertainty use theory to make testable predictions in a number of ways. One approach is to show that the observed strategy maximizes the geometric mean fitness instead of the arithmetic mean fitness (Boyce and Perrins, 1987; Philippi et al., 2001). The difficulty with such quantitative predictions is that they are often quite sensitive to errors in observed parameters of the model, such as the frequency of different kinds of environments. Instead, most studies test qualitative predictions. For example, if it can be shown

that a mixture of phenotypes⁵ performs better in the long term than any single phenotype, observed variation could be an adaptation to uncertainty (Saiah and Perrin, 1990; Menu et al., 2000; Evans et al., 2007). Stronger evidence is gained by comparing several populations or species with different amounts of uncertainty about the environment. If the amount of phenotypic diversity observed varies with the environmental risk as predicted by theory, that diversity is likely to be adaptive (Philippi, 1993; Clauss and Venable, 2000; Koops et al., 2003; Halkett et al., 2004; Venable, 2007).

The framework presented here can be used to make quantitative predictions about what phenotype or mixture of phenotypes would be best adapted to a particular set of conditions. In order to use it, a realistic, data-driven model of how different phenotypes fare in different environments is needed, and an accurate assessment of the frequency of different kinds of environments over the long term must be made. Then the procedure illustrated with the example of frog metamorphosis can be followed to predict the optimal response. While similar quantitative predictions could also be made without our framework—via numerical optimization or computer simulation—our graphical method of analysis provides a clearer picture of why bet-hedging may or may not be favored in any particular system. If bet-hedging is not adaptive, it could be because the region of strong trade-offs is so small that bet-hedging would almost never be adaptive. On the other hand, it might be because the environmental frequencies just happen to lie outside the appropriate range. Another reason for preferring our approach is that it gives a comprehensive picture of how sensitive the optimal strategy is to measurement errors in various parameters of the model. For example, small changes in the fitness functions can change strong trade-offs into weak trade-offs (compare Figures 3.3 and 3.5), going from a situation in which bet-hedging is often favored to a situation where it is never favored.

Our framework generates several novel qualitative predictions that could be tested using a comparative approach. The first is that the range of environmental uncertainty in which bet-hedging is adaptive is limited by the size of the region of strong trade-offs (see Figure 3.4.) Other models have indicated that bet-hedging is only adaptive when the variance in which phenotype is optimal exceeds a certain threshold value (Slatkin and Lande, 1976; Bull, 1987), suggesting that a minimal amount of environmental variance is necessary. Haccou and Iwasa (1995) note that this minimal variance threshold depends on how generalist the phenotypes are, and calculate it for several specific func-

⁵some particular mixture

tions describing the relationship between phenotype, environment and fitness. Our result is more general, because it is independent of the particular form of the fitness function. It is also more explicit, because it describes the environmental probability distributions themselves rather than simply the variance of those distributions. The second prediction is that the range of potentially optimal bet-hedging strategies is affected by the cost of plasticity. As long as there is no cost to plasticity, any given mixture of the right phenotypes can be optimal, for some particular amount of environmental uncertainty (see Figure 3.4, right panel.) On the other hand, if there is a cost to plasticity, some mixtures of the right phenotypes will never be optimal under any level of environmental uncertainty (see Figure 3.6, right panel.) Finally, we show that there is a linear relationship between the environmental frequencies and the optimal bet-hedging mixture of phenotypes. How sensitive the optimal mixture is to changes in environmental frequencies depends on how big the the range of potentially optimal bet-hedging strategies is, relative to the range of environmental frequencies in which bet-hedging is optimal. That means that if the region of strong trade-offs is small—say, making bet-hedging optimal only in the range from 50–52% dry years—even a tiny change in the environmental probabilities can have a big effect on the optimal mixture of phenotypes.

The idea that random phenotypic variation can lessen the negative impact of environmental stochasticity is often explained in terms of investment in the stock market. In a well-crafted stock portfolio, high-risk, high-return stocks are combined with low-risk, low-return stocks in a way that maximizes the expected rate of return. Similarly, bet-hedging genotypes often produce a combination of two life history strategies, one with high expected fitness, but high risk, and another with lower expected fitness and lower risk. Our framework builds upon this analogy. What we call a completely specialized phenotype is completely invested in one particular environment; in all other environments that individual is counted as a total loss. Any phenotype that is not completely specialized, however, is effectively invested in several different environments at once. Its proportional investment in different environments is defined by comparing the relative fitness of different types. Bet-hedging is a way of fine-tuning the total investment in different environments, by producing a variety of offspring whose average specialization level matches the probabilities of the different kinds of environments.

Two important factors in the evolution of specialists and generalists are left for future exploration. One is the effect of competition for resources, which can promote the evolution of within-

species polymorphism in resource use (see Skúlason and Smith, 1995, for a review). Examples of this are widespread in birds, amphibians, and fish and may arise through genetic divergence (e.g. Smith, 1993) or condition-dependent plasticity (e.g. Frankino and Pfennig, 2001). This mechanism for the generation of adaptive variation is driven by frequency-dependent selection, and can act independently of the kind of population-level environmental uncertainty we have considered here. As discussed in Section 3.4, our model cannot capture the full picture when there is frequency-dependent selection.⁶ However, recent work has extended the use of the fitness set into scenarios including frequency dependence (Rueffler et al., 2004; de Mazancourt and Dieckmann, 2004). Such an approach might be a useful extension for our model. Another exciting direction will be to consider predictive cues that can be used to direct phenotypic plasticity. For example, some tadpoles react to decreasing water levels in their pond by accelerating development (Denver et al., 1998). If the cues do not predict the environment perfectly, however, some environmental uncertainty may remain—making a combination of bet-hedging and plasticity a potentially useful strategy (DeWitt and Langerhans, 2004). As demonstrated in Chapter 4, our framework can easily and naturally be extended to explore the relationship between the amount of information and population-level uncertainty in a cue, and the optimal balance between predictive and stochastic plasticity.

Appendices

A A model of frog metamorphosis

Here we specify a simple model of frog metamorphosis. A tadpole is capable of metamorphosis after a minimum number days, t . After this point, it grows linearly with slope k . This determines the size of the frog at metamorphosis, $f(x) = k(x - t)$. However, the longer it spends as a tadpole, the more likely it is that the pond will dry up before it manages to metamorphose. In wet years, the time of drying for individual ponds is independent and distributed normally, with mean μ_w and standard deviation σ_w . In dry years, the time of drying for individual ponds is independent and distributed normally, with mean μ_d and standard deviation σ_d . The probability that a pond does not dry up by time x in dry years is one minus the cumulative distribution function for the appropriate

⁶or a combination of frequency- and density-dependent selection

normal distribution:

$$s_d(x) = 1 - \frac{1}{2} \left(\operatorname{erf} \left(\frac{x - \mu_d}{\sqrt{2}\sigma_d} \right) + 1 \right), \quad (\text{A-1})$$

where $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$. A similar expression holds for $s_w(x)$. The expected fitness for a tadpole metamorphosing after x days is $f_d(x) = k(x-t)s_d(x)$ in a dry year, or $f_w(x) = k(x-t)s_w(x)$ in a wet year. For Figures 3.2, 3.3, and 3.4, we use parameter values $t = 16$, $k = 1/12$, $\mu_d = 36$, $\sigma_d = 3$, $\mu_w = 56$, and $\sigma_w = 6$. For Figure 3.5 we increase the standard deviations to $\sigma_d = 6$ and $\sigma_w = 12$.

In Section 3.3.4, we add a cost of plasticity to the model. Compared to the same phenotype produced without plasticity, a plastically produced phenotype experiences a proportional decrease in fitness: $f'_d(x) = (1 - c_d)f_d(x)$ and $f'_w(x) = (1 - c_w)f_w(x)$. In Figure 3.6, we use $c_d = c_w = 0.05$. Because the cost is proportional to fitness, the extended fitness set created by all linear combinations of f'_d and f'_w , instead of f_d and f_w , shrinks but does not change shape. In addition, the extrapolation line used to calculate specialization levels is stretched without changing the proportions. This property of a fitness-proportional cost function makes it easier to compare the results with and without cost. However, any other kind of cost function could be used, as long as it is constant with respect to the frequency of phenotypes that are produced.

B Calculating specialization levels

Given a set of phenotypes (as defined by the region of strong trade-offs, see Section 3.3.1) we would like to write the fitnesses of any mixture in each environment as a mixture of completely specialized phenotypes, having non-zero fitness in only one environment. This corresponds to writing the fitness f_{xk} of phenotype x in environment k as a linear combination of values d_k , each of which represents the fitness of a phenotype completely specialized in environment k . Each coefficient in these linear combinations, s_{xk} , is the specialization level of phenotype x in environment k .⁷ The problem can be formulated as a matrix equation:

$$\mathbf{F} = \mathbf{SD}, \quad (\text{A-2})$$

⁷The specialization level of a phenotype x in some environment k can be thought of as the proportion of the phenotype's reproductive effort that is invested in environment k .

where \mathbf{S} is a stochastic matrix with rows summing to 1, and \mathbf{D} is a diagonal matrix.⁸

We assume that \mathbf{F} is a square matrix, that is, there are as many phenotypes as environments. This will be true as long as there are strong fitness trade-offs between specializing in all possible environments. In this case (barring singularities) both \mathbf{F} and \mathbf{D} are invertible, so

$$\mathbf{D}^{-1} = \mathbf{F}^{-1}\mathbf{S}. \quad (\text{A-3})$$

Multiplying on the right by a column vector of ones, we get the row sums

$$\text{row sum}(\mathbf{D}^{-1}) = \text{row sum}(\mathbf{F}^{-1}) \quad (\text{A-4})$$

because a stochastic matrix has all row sums equal to 1. This defines the diagonal matrix \mathbf{D} completely, allowing the calculation of the specialization in each environment as

$$\mathbf{S} = \mathbf{F}\mathbf{D}^{-1}. \quad (\text{A-5})$$

This method will yield positive specialization levels as long as the diagonal matrix \mathbf{D} has only positive entries. That is, returning to the graphical viewpoint illustrated for two dimensions in Figure 3.4, the hyperplane passing through all phenotypes plotted in fitness space must intersect each axis at a positive point. This will be true as long as there is a trade-off between fitness in all different environments, as we have assumed.

⁸Here, the specialization matrix \mathbf{S} and the environmental productivity matrix \mathbf{D} are the matrices we would like to find, given the fitness matrix \mathbf{F} .

Chapter 4

DEVELOPMENTAL PLASTICITY AND ENVIRONMENTAL INFORMATION: HOW THE VALUE OF INFORMATION DEPENDS ON ITS STRUCTURE

4.1 Introduction

Organisms living in variable environments face a problem of prediction. The better they can predict and respond to fitness-relevant changes in the environment, the better the chance they will have to survive and reproduce. Developmental polyphenism, which creates multiple discrete phenotypes from the same genotype, is an important adaptation in this context (Levins, 1968; Stearns, 1989; Moran, 1992). Non-genetic variation in phenotypes is commonly modeled as stemming from a heritable developmental program—a branching network of developmental trajectories, with switches that determine which branch will be followed. These switches may be regulated conditionally on external cues received during development, or they may be regulated stochastically, for example via internally generated developmental noise (West-Eberhard, 2003).

How does selection in a variable environment influence the balance between stochastically and conditionally generated phenotypic diversity? If the information available in the developmental environment is sufficient to reliably predict the optimal adult phenotype, organisms may benefit from harnessing that information to deterministically guide the developmental trajectory (Cohen, 1967; Via and Lande, 1985; Lively, 1986a; Moran, 1992). This kind of conditional regulation is often called adaptive developmental plasticity.

However, if there is no predictive information in the developmental environment, even randomly generated phenotypic variation can be adaptive (Cohen, 1966). Phenotypic diversity within a genetic lineage helps to mitigate the drastic effects of events like drought or a late spring freeze, which could otherwise wipe out the entire lineage at once. Stochastic regulation as an adaptation to population-level risk is known as bet-hedging or “adaptive coin-flipping” (Cooper and Kaplan, 1982; Seger and Brockmann, 1987). While modeling efforts have often focused on these extreme cases, in the real world there will often be some information in the developmental environment, but not enough for

perfect prediction. These conditions can select for a combination of stochastic and conditional regulation, yielding phenotypes that are only partially determined by their developmental environments (DeWitt and Langerhans, 2004).

In this paper, we explore the adaptive significance of developmental cues that give partial information about an environment that changes from one generation to the next. In particular, we focus on the way that error in the cue is distributed between individuals: is the error in a cue common to all individuals observing the cue, or are the errors in cues independent from one individual to the next? For example, tadpoles living in ephemeral ponds seem to be able to accelerate their development as the water in the pond dries up (Newman, 1992). However, it is still unknown whether the tadpoles react directly to changes in the water depth, or to other correlated cues like the temperature of the water or concentration of compounds (Denver et al., 1998). These different types of cues may differ, both in how well they predict whether the pond will dry up, and in how reliably individuals can observe them. Previous theoretical work has shown that the adaptive value of a developmental cue—when any error in the cue affects the entire population—depends on the amount of information it contains (Bergstrom and Lachmann, 2004). How does the distribution of error between individuals, in combination with the amount of error, affect the developmental response to different kinds of cues?

4.2 A model of developmental plasticity in response to unreliable cues

Consider a population of individuals in a fluctuating environment. In each generation, the state of the environment e is drawn from a distribution $\Pr(e)$ over a discrete set of possible environments. We consider population-level risk, in which the environmental state affects every individual in the population according to its phenotype x , via the fitness function $f(x, e)$. In this model, we focus particularly on discrete phenotypes, as would be produced by developmental polyphenism. If there is no frequency-dependent selection, a good predictor of a genotype's success in the long term is the geometric mean fitness (Dempster, 1955; Cohen, 1966), or alternatively its expected long-term growth rate (the log of the geometric mean.) This depends not just on the average fitness of individuals of that genotype, but also how their fitness varies within and between generations (Frank and Slatkin, 1990).

Fitness variation between individuals is a result of both variation in the selective environment, and variation in phenotype. In our model, the environment varies at the population level, so it can only contribute to variation in fitness between generations. Phenotypic diversity, on the other hand, can contribute to variation in fitness both within and between generations. The phenotypic diversity generated by a particular genotype depends on the developmental strategy it encodes. If that strategy depends on developmental cues, the structure of variation in those cues also contributes to phenotypic variation. Below, I describe how the developmental strategy and the error in the developmental cue combine to create phenotypic diversity within and between generations, and how this phenotypic diversity then affects the long-term growth rate achieved by a particular developmental strategy.

4.2.1 *Types of developmental strategies*

An individual's phenotype is determined according to its developmental strategy. Each individual receives a developmental cue c , and in response develops a single phenotype x . A developmental strategy is represented as a row stochastic matrix \mathbf{G} , where each entry $g_{cx} = g(x|c)$ represents the probability of developing into the phenotype x , conditional on receiving a developmental cue c . Several special cases deserve particular notice:

- a) A *pure strategy* always produces a single phenotype, regardless of the developmental environment, e.g. $g(x_1) = 1$.
- b) A *conditional strategy* produces different phenotypes in response to different developmental cues, e.g. $g(x_1|c_1) = 1$ and $g(x_2|c_2) = 1$.
- c) A *stochastic strategy* randomly produces one of several phenotypes, in certain proportions, regardless of developmental environment. This corresponds to the case when all rows of the matrix \mathbf{G} are equal. We will represent this kind of strategy by a stochastic vector \vec{g} , with entries $g_x = g(x)$.
- d) A *conditional stochastic strategy* can produce any of the possible phenotypes x in response to any of the developmental cues c —only the proportions of the phenotypes vary. This is the

Table 4.1: Summary of mathematical notation.

Matrix or vector	Entries	Description
environment		
\vec{p}	$p_e = \Pr(e)$	Probability of each environment
\mathbf{Q}	$q_{qe} = \Pr(e, q)$	Joint probability of each environment and population-level cue
$\tilde{\mathbf{Q}}$	$\tilde{q}_{qe} = \Pr(e q)$	Conditional probability of each environment, given the population-level cue
\mathbf{C}	$c_{qc} = \Pr(c q)$	Conditional probability of each developmental cue, given the population-level cue
fitness		
\mathbf{F}	$f_{xe} = f(x, e) = s(e x)d(e)$	Fitness of each phenotype in each environment
\mathbf{S}	$s_{xe} = s(e x)$	Proportion of reproductive effort invested by a phenotype in each environment
\mathbf{D}	$d_{ee} = d(e), d_{ee'} = 0$	Productivity of each environment, or the fitness that would be achieved by a phenotype that invested all of its reproductive effort in that environment
developmental strategy		
\vec{g}	$g_x = g(x)$	Probability of each phenotype, as produced by a stochastic strategy
\mathbf{G}	$g_{cx} = g(x c)$	Conditional probability of each phenotype given the developmental cue, as produced by a conditionally stochastic strategy

most general type of strategy, where all entries of the matrix \mathbf{G} are free to vary.

A stochastic strategy (or a conditional stochastic strategy) will be called *fully stochastic* if all entries of the matrix \mathbf{G} are non-zero—meaning, it is always possible to develop into any one of the phenotypes, no matter the developmental cue. If some entries are zero, the strategy is only partially stochastic; in Figure 4.1, these are strategies that always produce one phenotype in response to one cue, but produce a mixture of phenotypes in response to the other.

The measures of information theory provide a way to quantify uncertainty—as the entropy of a random variable—and information—as a reduction in entropy (Cover and Thomas, 1991). The phenotypic diversity that is created by a developmental strategy can thus be quantified in terms of the uncertainty about which phenotype will be produced:

$$\begin{array}{l} \mathbf{total} \\ \mathbf{phenotypic} \\ \mathbf{diversity} \end{array} \qquad H(X) = - \sum_x g(x) \log g(x). \qquad (4.1)$$

Two sources of variation are used to generate this phenotypic diversity. How much variation in phenotype is explained by variation in developmental cues? This is represented by the Shannon information between the cue and the phenotype:

$$\begin{array}{l} \mathbf{conditional} \\ \mathbf{phenotypic} \\ \mathbf{diversity} \end{array} \qquad I(X;C) = \sum_c \Pr(c) \sum_x g(x|c) \log \frac{g(x|c)}{g(x)}. \qquad (4.2)$$

How much variation in phenotype is explained by stochastic variation during development? This can be represented by the conditional entropy of the phenotype, given the developmental cue:

$$\begin{array}{l} \mathbf{stochastic} \\ \mathbf{phenotypic} \\ \mathbf{diversity} \end{array} \qquad H(X|C) = - \sum_c \Pr(c) \sum_x g(x|c) \log g(x|c), \qquad (4.3)$$

which is simply how much uncertainty remains about the phenotype, once the cue is known. The sum of these two quantities, the conditional variation and the stochastic variation, is equal to the total phenotypic variation: $I(X;C) + H(X|C) = H(X)$. The different types of developmental strategies, and the balance of stochastic and conditional regulation in each, are illustrated in Figure 4.1 for the simplest case of two cues and two phenotypes.

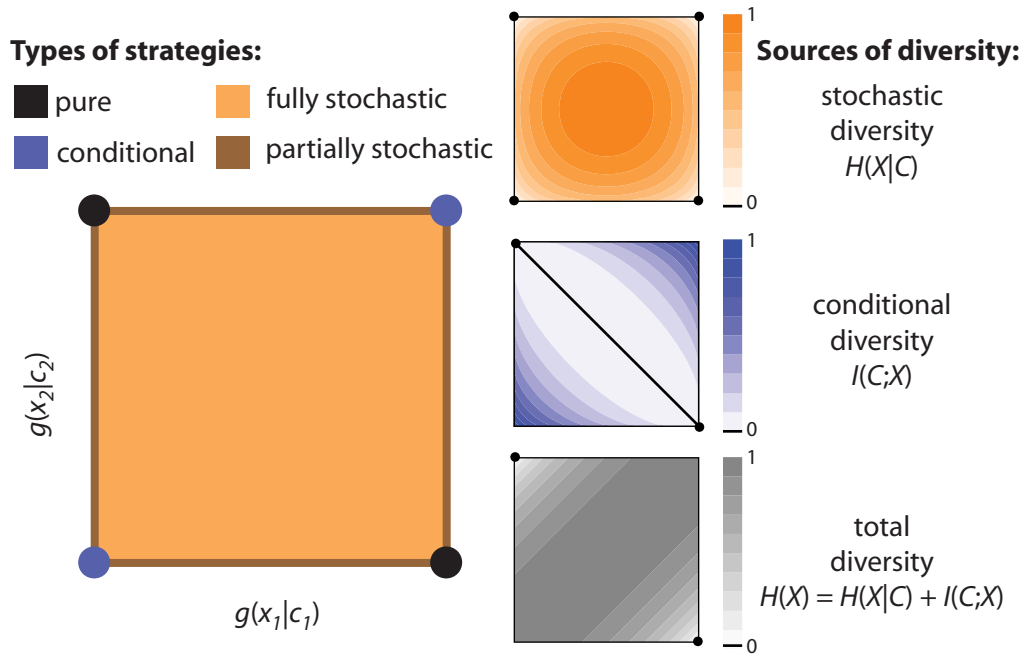


Figure 4.1: Types of developmental strategies. In the simplest case, there are just two cues, c_1 and c_2 , and two phenotypes, x_1 and x_2 . Each strategy is defined by a pair of conditional probabilities, $g(x_1|c_1)$ and $g(x_2|c_2)$. There are just four different kinds of strategies: (1) a pure strategy, which always produces the same phenotype; (2) a conditional strategy, which produces a different phenotype for each cue; (3) a fully stochastic strategy, which produces a mixture of phenotypes for both cues; and (4) a partially stochastic strategy, which produces a mixture of phenotypes for one cue, and produces just one phenotype for the other. At left, all possible developmental strategies are classified into one of these four types. At right, these strategies are further characterized according to the total amount of phenotypic diversity they produce and how that diversity is regulated. At top, the amount of diversity produced as a result of stochastic regulation is represented according to the conditional uncertainty in the phenotype given the cue, $H(X|C)$. In the middle, how much diversity is produced by conditioning on the cue is represented according to the information in the cue about the phenotype, $I(C;X)$. At bottom, the total amount of phenotypic diversity as produced by both mechanisms is represented according to the uncertainty about the phenotype, $H(X)$. In all cases, higher phenotypic diversity is indicated by more intense colors, while zero diversity is specially indicated in black.

4.2.2 *Types of error in a developmental cue*

Stochastic regulation of development creates phenotypic diversity within a generation. However, conditional regulation of development can produce phenotypic variation both within and between generations, depending on the way that the developmental cue itself varies. How is the developmental cue received by one individual related to the cue received by others in the same generation? If there is any information in the cue about the environmental state—which we assume is common to everyone—there must be some correlation between individual cues. How much more correlation is there between individual cues, beyond what is created by common information about the environment?

At one extreme, if every individual in the same generation receives the same developmental cue, any error in the cue affects the entire population at once. This kind of cue creates error at the population level, like the environmental uncertainty itself, and can thus drive the evolution of conditional bet-hedging. The optimal developmental strategy bet-hedges in response to each cue, according to how much uncertainty is left about the state of the environment (Cohen, 1967; Haccou and Iwasa, 1995). At the other extreme, if each individual receives a cue independently given the state of the environment, all of the correlation between developmental cues is a result of information about the environment. The error in such cues is purely at the individual level, and limits the individual's ability to match its phenotype to the selective environment.

Most developmental cues probably lie somewhere between the two extremes, their error being neither completely at the individual level nor completely at the population level. Mathematically, we can describe the situation quite generally by specifying a joint distribution for the developmental cues of all individuals within a generation, conditional on the state of the environment: $\Pr(c_1, c_2, \dots, c_n | e)$. It is often possible, however, to think of a developmental cue in the following terms. All individuals attempt to observe some aspect of the global environment, q , which is correlated with the environment e . However, each individual i makes some error in observation, leading to the individual cue c_i . These individual cues are independent and identically distributed, given the population-level cue q . That means we can rewrite the joint probability distribution as follows: $\Pr(c_1, c_2, \dots, c_n | e) = \sum_q \Pr(q | e) \prod_i \Pr(c_i | q)$ (see Appendix A for details.)

The power of this approach is that it allows us to identify how much of the phenotypic variation

generated by a conditionally stochastic strategy occurs between generations. The phenotypic variation produced within any generation depends on the distribution of cues in that generation, which is defined by the population-level variable q : $\bar{g}(x|q) = \sum_c \Pr(c|q)g(x|c)$. Using a row stochastic matrix \mathbf{C} to represent the conditional probability distribution $\Pr(c|q)$, the phenotypic variation in response to the population-level cue is more succinctly represented as \mathbf{CG} . The total phenotypic diversity $H(X)$ can now be divided up into a within-generation component $H(X|Q)$ and a between-generation component $I(X;Q)$. The within-generation diversity is generated by stochastic developmental regulation and conditional response to individual-level variation in developmental cues, while the between-generation diversity is generated only by conditional response to population-level variation in developmental cues.

4.2.3 *Individual fitness and the growth rate of a lineage*

To find the optimal developmental strategy, we would like to characterize each strategy in terms of its expected long-term growth rate. How does the variation in phenotype created by the developmental strategy translate into individual variation in fitness? How does variation in individual fitness further translate into the long-term growth rate of the lineage?

An individual with phenotype x , in environment e , has fitness specified by the fitness matrix \mathbf{F} . In a single generation, a lineage grows according to its average fitness within that generation. This average within-generation fitness depends on the phenotypic diversity generated by the strategy, and on the selective environment. The long-term growth rate is the average, over many generations, of the log average fitness within each generation. The variance in fitness within generations does not affect the growth rate, as long as the number of individuals in the lineage is large enough that the observed average fitness reliably reflects the mean. However, the variance in the average fitness per generation matters, with higher variance from one generation to the next tending to slow the growth rate.

First consider a developmental strategy \vec{g} that uses only stochastic regulation. The amount of phenotypic diversity does not vary from one generation to the next, so the mean fitness within a generation depends only on the environmental state e : $\bar{f}(\vec{g}|e) = \sum_x g(x)f(x, e)$. In matrix notation, this can be written as a vector over environmental states $\vec{g}\mathbf{F}$. The variation between generations thus

depends on the probability of different environmental states $\Pr(e)$. The expected long-term growth rate of a purely stochastic strategy is then conveniently expressed in matrix notation as follows:

$$\begin{array}{l} \text{growth rate,} \\ \text{without cue} \end{array} \quad r(\vec{g}) = \vec{p} \cdot \log \vec{g}\mathbf{F}, \quad (4.4)$$

where the operator \cdot represents the dot product of two vectors: $\vec{a} \cdot \vec{b} = \sum_i a_i b_i$.

Next, consider a developmental strategy \mathbf{G} that uses a combination of stochastic and conditional regulation. The phenotypic diversity in each generation depends on the distribution of developmental cues. The mean fitness within a generation thus depends both on the environmental state e and the population-level distribution of cues defined by q : $\bar{f}(\mathbf{G}|q, e) = \sum_c \Pr(c|q) \sum_x g(x|c) f(x, e)$. This can be written as a matrix with a row for each population-level cue, and a column for each environmental state, \mathbf{CGF} . The expected long-term growth rate then depends on the probability of different types of generations, as defined by the combination of the environmental state and population-level cue: $\sum_{q,e} \Pr(q, e) \log \bar{f}(\mathbf{G}|q, e)$. We will represent the probability of different types of generations with the matrix \mathbf{Q} , with entries $q_{qe} = \Pr(e, q)$. This yields the following simple representation for the growth rate of a conditional stochastic strategy:

$$\begin{array}{l} \text{growth rate,} \\ \text{general cue} \end{array} \quad r_B(\mathbf{G}) = \mathbf{Q} : \log \mathbf{CGF}. \quad (4.5)$$

The operator $:$, known as the Frobenius inner product, is like a dot product for matrices: $\mathbf{A} : \mathbf{B} = \sum_{i,j} A_{ij} B_{ij}$. This is just like Equation 4.4, but with matrices over (q, e) instead of vectors over e .

4.3 Optimal developmental strategies and the value of a cue

How does a developmental strategy reflect features of the environment to which it is adapted? When do we expect organisms to use predominantly stochastic development, and when predominantly conditional regulation? When will we see a mixture of both? To answer these questions, we characterize the optimal developmental strategy in a fluctuating environment, where developmental cues contain error at both the population level and the individual level. Furthermore, we would like to know how the adaptive value of a developmental cue depends on its distribution of error. To do this, we compare the optimal growth rate using a developmental cue to the optimal growth rate with no cue. The difference between these two growth rates is the adaptive value of using the cue (Bergstrom and

Lachmann, 2004). Over the long term, we may be likely to observe developmental strategies that utilize the cues with the highest adaptive value.

To find the optimal strategy in each case, we will use the machinery developed in Chapter 3 for finding the optimal bet-hedging strategy when no information is available. The primary insight is that the optimal developmental strategy invests in each possible environment according to its probability of occurring. We can define a *specialization profile* for each phenotype, which indicates how much of its reproductive effort that phenotype allocates to each environment. To do this, we rewrite the fitness matrix as a composition of two matrices, $\mathbf{F} = \mathbf{SD}$. In this breakdown, \mathbf{S} is a row stochastic matrix with individual entries $s_{xe} = s(e|x)$ representing how specialized phenotype x is in environment e ; \mathbf{D} is a diagonal matrix with entries $d_{ee} = d(e)$ representing how productive each environment is overall. The fitness of a phenotype x in environment e is a product of how specialized it is in that environment and the productivity of that environment: $f(x, e) = s(e|x)d(e)$. The specialization $s(e|x)$ of a phenotype x in environment e is the proportion of its reproductive effort the phenotype invests in that environment. The productivity of the environment, $d(e)$, is the maximum amount of fitness that could be achieved in that environment, if a phenotype were to be completely specialized in that environment. The optimal stochastic strategy, \vec{g}^* , is the one that makes the average specialization profile match the environmental probability distribution:

$$\begin{array}{l} \text{optimal strategy,} \\ \text{without cue} \end{array} \quad \vec{g}^* \mathbf{S} = \vec{p}. \quad (4.6)$$

Whether this is feasible depends on whether the optimal specialization profile, \vec{p} , lies within the achievable range defined by the specialization matrix \mathbf{S} .

The long-term growth rate achieved with the optimal strategy, within the range where it can be attained, is obtained by combining Equations 4.4 and 4.6:

$$\begin{array}{l} \text{optimal} \\ \text{growth rate,} \\ \text{without cue} \end{array} \quad r(\vec{g}^*) = \vec{p} \cdot \log \vec{g}^* \mathbf{SD} = \vec{p} \cdot \log \vec{p} \mathbf{D}. \quad (4.7)$$

If the optimal specialization level cannot be achieved using a fully stochastic strategy, the best-fitting partially stochastic strategy will be used instead. The optimal strategy is the one whose specialization profile comes closest to matching the environmental probabilities, measured in terms of the Kullback-Leibler divergence between probability distributions (see Chapter 3.)

4.3.1 Developmental cues with population-level error

Previous models of the evolution of plasticity in response to population-level risk have only considered cues with errors at the population level (e.g. Cohen, 1967; Moran, 1992; Haccou and Iwasa, 1995). The organism is assumed to observe some feature of the global environment during development, and use this to predict some feature of the future selective environment. The error in the cue is a result of imperfect correlation between the state of the environment at the time of development and the future state of the environment. The results presented in this section are already known; we review them here because they provide a good basis for comparison for the rest of the results.

The optimal response to a developmental cue that is received by all individuals in the population is particularly easy to calculate. The reason for this is that the response to any specific developmental cue contributes to the growth rate only in generations where that cue is received, and the average long-term growth rate is simply an average of the growth rate achieved for each different developmental cue. Using the notation adopted above for Equation 4.5, the special case of no individual-level error corresponds to $c = q$, meaning that \mathbf{C} is the identity matrix \mathbf{I} . Equation 4.5 can therefore be rewritten as follows:

$$\begin{aligned}
 \text{growth rate,} & & r_P(\mathbf{G}) &= \mathbf{Q} : \log \mathbf{G}\mathbf{F} \\
 \text{population cue} & & &= \sum_c \Pr(c) \vec{q}_c \cdot \log \vec{g}_c \mathbf{F}, \quad (4.8)
 \end{aligned}$$

where \vec{q}_c is a stochastic vector representing the conditional probability distribution of environments $\Pr(e|c)$. The response to each possible developmental cue c , represented by \vec{g}_c , can therefore be optimized independently of the others, according to Equation 4.6: $\vec{g}_c^* \mathbf{S} = \vec{q}_c$. Let the matrix $\tilde{\mathbf{Q}}$ be a conditional probability matrix whose rows are the vectors \vec{q}_c . Then the optimal developmental strategy \mathbf{G}^* can be written in matrix notation as follows:

$$\begin{aligned}
 \text{optimal strategy,} & & \mathbf{G}^* \mathbf{S} &= \tilde{\mathbf{Q}}. \quad (4.9) \\
 \text{population cue} & & &
 \end{aligned}$$

This optimal strategy is achievable if the optimal specialization profile in response to each cue, \vec{q}_c , lies within the range defined by the specialization matrix \mathbf{S} .

For example, consider the simplest case of two environments, two cues, and two phenotypes. As long as the conditional probability $\Pr(e_1|c_1)$ is between $s(e_1|x_1)$ and $s(e_1|x_2)$, the optimal strategy is to bet-hedge in response to c_1 so that the average specialization in e_1 matches the conditional probability of e_1 , that is, $g(x_1|c_1)s(e_1|x_1) + g(x_2|c_1)s(e_1|x_2) = \Pr(e_1|c_1)$. Similarly, as long as the conditional probability $\Pr(e_2|c_2)$ is between $s(e_2|x_1)$ and $s(e_2|x_2)$, the optimal strategy is to bet-hedge in response to c_2 so that the average specialization in e_2 matches the conditional probability of e_2 : $g(x_1|c_2)s(e_2|x_1) + g(x_2|c_2)s(e_2|x_2) = \Pr(e_2|c_2)$. This creates a square region (central, in orange) in which a fully stochastic strategy is optimal, as illustrated in Figure 4.2 for the specialization matrix $\mathbf{S} = \begin{pmatrix} 0.70 & 0.3 \\ 0.20 & 0.8 \end{pmatrix}$.

The optimal growth rate for a developmental strategy using a cue with no individual-level error can now be found by substituting the optimal strategy from Equation 4.9 into the growth rate given by Equation 4.8:

$$\begin{array}{l} \text{optimal} \\ \text{growth rate,} \\ \text{population cue} \end{array} \quad r_P(\mathbf{G}^*) = \mathbf{Q} : \log \mathbf{G}^* \mathbf{S} \mathbf{D} = \mathbf{Q} : \log \tilde{\mathbf{Q}} \mathbf{D}. \quad (4.10)$$

In comparison, the growth rate under the optimal stochastic strategy without developmental cues is given by Equation 4.7. The adaptive value of a population-level cue is the difference between these two growth rates:

$$\begin{array}{l} \text{adaptive value,} \\ \text{population cue} \end{array} \quad \begin{aligned} V_P(C) &= \mathbf{Q} : \log \tilde{\mathbf{Q}} - \vec{p} \cdot \log \vec{p} \\ &= H(E) - H(E|C) = I(E;C), \end{aligned} \quad (4.11)$$

which is simply the Shannon information between the cue and the environment. This result holds precisely as long as the optimal specialization levels can be attained, leading to a fully stochastic strategy in response to each cue. Otherwise, the information is an upper limit for the value of the cue (Bergstrom and Lachmann, 2004). In Figure 4.3, we illustrate the relationship between the value of the cue and the information it contains for the example used in Figure 4.2. In the central region, where fully stochastic strategies are optimal, the value is equal to the information, while elsewhere it is strictly less.

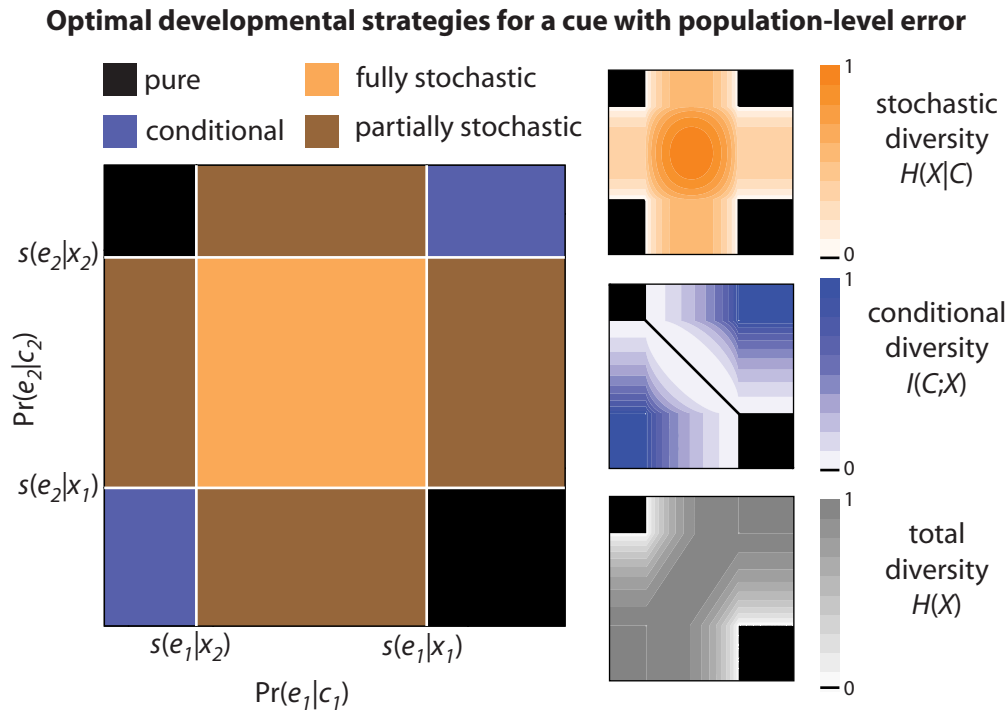


Figure 4.2: Optimal developmental strategies in response to a cue with no individual-level error, and varying amounts of population-level error. We illustrate how the optimal developmental strategy, $g(x|c)$, depends on the reliability of the cue $Pr(e|c)$, in the simple case of two environments, e_1 and e_2 , two cues c_1 and c_2 , and two phenotypes, x_1 and x_2 . At left, the optimal strategy for each cue-environment relationship is classified according to its type: pure, conditional, fully or partially stochastic, with colors as in Figure 4.1. A fully stochastic strategy is optimal whenever the conditional probability of each environment lies within the range defined by the specialization matrix—here, we use $s(e_1|x_1) = 0.7$ and $s(e_2|x_2) = 0.8$. In the column at right, the optimal strategy is characterized according to how much phenotypic diversity it produces, through stochastic (top) and conditional (middle) regulatory mechanisms and in total (bottom.) As in Figure 4.1, more intense colors indicate greater phenotypic diversity. Notice the large regions of zero diversity, shown in black.

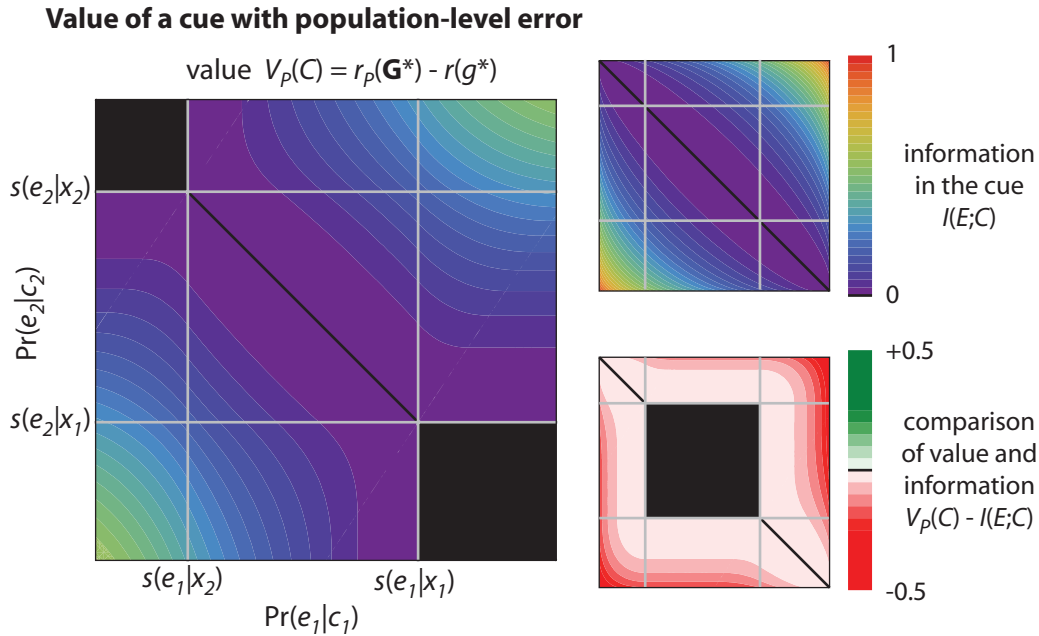


Figure 4.3: The value of a cue with no individual-level error, and varying amounts of population-level error. The value of the cue, $V_P(C)$, is equal to the difference in the optimal growth rates of two lineages: one that uses a cue with population-level error, and one that uses no cue at all. To illustrate the relationship of the amount of information a cue with population-level error contains about the environment to the value it confers, we return to the example of Figure 4.2, where $s(e_1|x_1) = 0.7$ and $s(e_2|x_2) = 0.8$. At left is a contour plot of the value of the cue, as its reliability varies. For comparison, the amount of information in the cue, as measured by the Shannon information $I(E;C)$, is plotted on the upper right, using the same color scheme as the value. On the lower right is the difference between the two. If the cue does not contain too much information—enough uncertainty remains that bet-hedging is still worthwhile—the increase in growth rate made possible by using the cue is exactly the amount of information in the cue. Otherwise, the cue is always worth less than the amount of information it contains. To make the correspondence to Figure 4.2 clearer, the grid dividing the regions where different types of strategies are optimal is superimposed in gray.

4.3.2 *Developmental cues with individual-level error*

In the last section we considered individuals that can perfectly observe some feature of the environment that is only indirectly related to which phenotype will have the highest fitness. Here we consider the opposite: individuals who can imperfectly observe the fitness-relevant state of the environment itself. That is, the cue itself is perfectly reliable, but individuals make mistakes in observation. We refer to this kind of error as individual-level error.

What is the optimal developmental response to cues with individual-level error? As described in Section 4.2.3, what is important for the growth rate is the variation in phenotypic diversity that the strategy produces, within and between generations. If there is no population-level error in the cue, the distribution of cues in each generation depends only on the state of the environment. The average fitness in different kinds of generations is expressed in Equation 4.5 as \mathbf{CGF} . However, this matrix takes into account all possible combinations of the population-level cue q and the environmental state e . In the special case of no population-level error, the population-level cue q is identical to the environmental state e . That means only the diagonal elements of \mathbf{CGF} contribute to the growth rate:

$$\begin{array}{l} \text{growth rate,} \\ \text{individual cue} \end{array} \quad r_I(\mathbf{G}) = \vec{p} \cdot \log \text{diag}(\mathbf{CGF}). \quad (4.12)$$

This greatly reduces the dimensionality of the optimization problem.

To illustrate the impact of this reduced dimensionality, we return to the usual case of two environments, two cues, and two phenotypes. Until now, we have described a conditional stochastic strategy as choosing a different “developmental coin” to flip, depending on the developmental cue observed. However, any such strategy can also be described as choosing a conditional strategy, depending on the outcome of a developmental coin flip. For example, the developmental strategy represented by the matrix $\mathbf{G} = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}$ can alternatively be expressed as follows: $0.3 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 0.4 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0.3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. According to this representation, the developmental strategy throws a three-sided coin with probabilities (30%, 40%, 30%), and then chooses one of three strategies. These three strategies are to always develop phenotype x_1 , regardless of the conditions, always develop phenotype x_2 regardless of the conditions, or develop x_1 in response to c_1 and x_2 in response to c_2 . In fact, every strategy can be represented as a linear combination of these three, or as a combination of the two pure strategies

plus the opposite conditional strategy. In this view, these four strategies—the corners of Figure 4.1—can be treated like phenotypes. The problem is then simply to find the optimal bet-hedging fractions of those phenotypes, given some level of uncertainty about the two possible environments—as described in Chapter 3. However, when only two different environments are possible, a bet-hedging strategy need incorporate at most two different phenotypes. In this case, those two phenotypes will be a plastic strategy and a pure strategy. That means a strictly optimal strategy will bet-hedge in response to at most one of the two cues; a fully stochastic strategy is never optimal (see Figure 4.4.)

The biggest difference between a cue with only individual-level error and one with only population-level error is that for the former, fully stochastic development is never strictly optimal. Even cues that contain almost no information about the environment can induce a strong conditional response: compare the points lying near the anti-diagonal line $\Pr(e_1|c_1) + \Pr(e_2|c_2) = 1$ in Figures 4.2 and 4.4. At the extreme, a completely uninformative cue can be used as the only source of variation, if the variation produced that way matches the optimal level of bet-hedging without a cue. For example, if $\Pr(c_1)s(e_1|x_1) + \Pr(c_2)s(e_1|x_2) = \Pr(e_1)$, then producing x_1 in response to c_1 and x_2 in response to c_2 will create just the right level of specialization in environment e_1 (Figure 4.4, point shown in white.)

Which has greater adaptive value, a cue with population-level error, or another cue with the same amount of error but at the individual level? First of all, we compare the growth rates of a single developmental strategy, responding to the two different kinds of cues. Writing out Equations 4.8 and 4.12 in full form,

$$\begin{aligned} r_P(\mathbf{G}) &= \sum_e \Pr(e) \sum_c \Pr(c|e) \log \sum_x g(x|c) f(x, e) \\ r_I(\mathbf{G}) &= \sum_e \Pr(e) \log \sum_c \Pr(c|e) \sum_x g(x|c) f(x, e), \end{aligned}$$

we see that the only difference between the two is the relative placement of the log and the average over individual cues. Jensen's inequality therefore implies that $r_I(\mathbf{G}) \geq r_P(\mathbf{G})$: the growth rate under individual-level error is never lower, no matter what strategy is being used. The *optimal* growth rate for a cue with individual-level error is therefore always at least as great as the optimal growth rate for the corresponding population-level error. The value of a cue with individual-level error is illustrated in Figure 4.5; comparison with Figure 4.3 shows that the value of the cue with individual-level

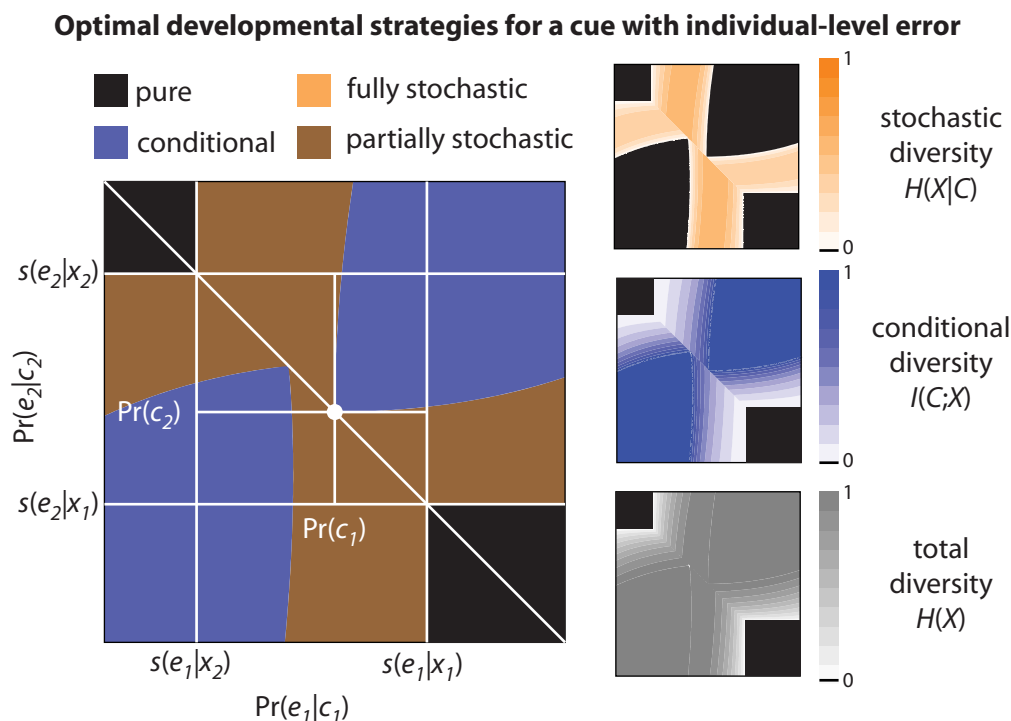


Figure 4.4: Optimal developmental strategies in response to a cue with varying amounts of individual-level error, and no population-level error. Just as we did for cues with population-level error (Figure 4.2) we show the relationship of the optimal developmental strategy $g(x|c)$ to the reliability of the cue, $Pr(e|c)$. To make the comparison straightforward, we use the same specialization matrix \mathbf{S} , with $s(e_1|x_1) = 0.7$ and $s(e_2|x_2) = 0.8$. All that differs is the way that errors in the developmental cues c are distributed. As in Figure 4.2, we categorize the optimal developmental strategies according to type in the left panel. Note that the conditions for a pure strategy to be optimal (in black) are independent of the level of error. However, the conditions for a conditional strategy to be optimal (in blue) are greatly expanded, while a fully stochastic strategy is never strictly optimal. Along the center diagonal line, the cues are completely uninformative; no single strategy is optimal, because all strategies that produce the right overall levels of diversity are equivalent. In the column at right, the optimal strategy is characterized according to how much phenotypic diversity it produces, through stochastic (top) and conditional (middle) regulatory mechanisms and in total (bottom.) As in Figure 4.1, more intense colors indicate greater phenotypic diversity. Notice the large regions of zero diversity, shown in black.

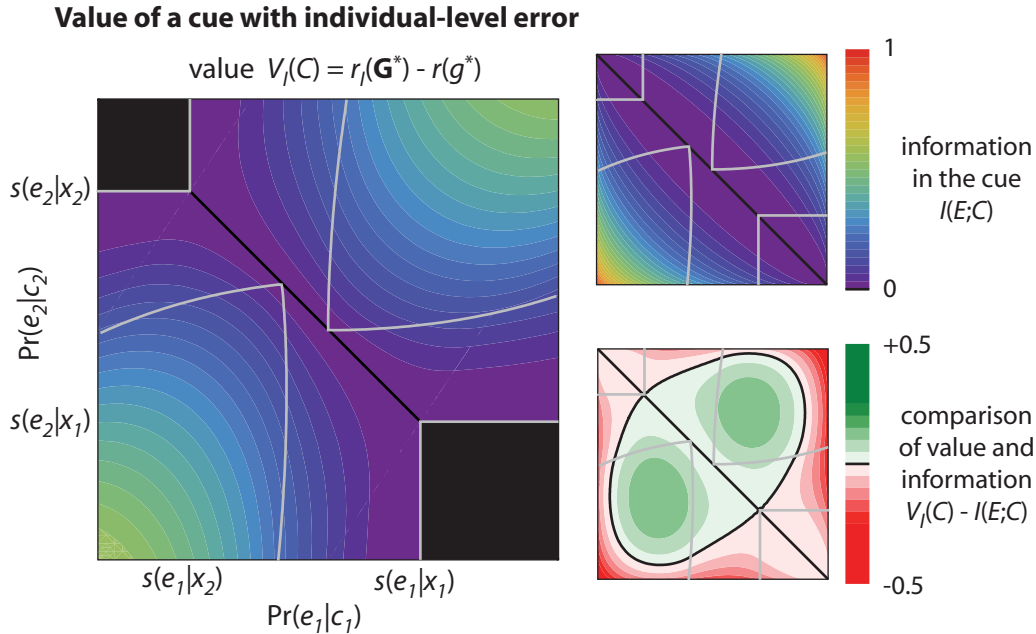


Figure 4.5: The value of a cue with varying amounts of individual-level error, and no population-level error. The value of the cue, $V_I(C)$, is equal to the difference in the optimal growth rates of two lineages: one that uses the cue with error at the individual level, and one that uses no cue at all. To illustrate the relationship of the amount of information a cue with individual-level error contains about the environment to the value it confers, we return to the example of Figure 4.4, where $s(e_1|x_1) = 0.7$ and $s(e_2|x_2) = 0.8$. At left is a contour plot of the value of the cue, as its reliability varies. For comparison, the amount of information in the cue, as measured by the Shannon information $I(E;C)$, is plotted on the upper right, using the same color scheme as the value. On the lower right is the difference between the two. For a wide range of cues, the value of the cue is in fact greater than the amount of information it contains about the environment (in green; compare to Figure 4.3.) The regions where different types of strategies are optimal, as calculated in Figure 4.4, are outlined in gray.

error is always greater, except where a pure strategy is optimal (in black), where the cue contains no information (along the diagonal line) and where the cue contains no error (opposite corners, in green.) Furthermore, comparison to the Shannon information in the cue about the environment shows that, unlike a cue with population-level error, a cue with individual-level error can sometimes have a value that is even greater than its information content (Figure 4.5, lower right.)

4.3.3 *Developmental cues with error at both levels*

The kinds of developmental cues considered in the previous two sections had error either at the individual level, which could stem from an inability to precisely observe relevant features of the environment during development, or error at the population level, which could be a result of truly unpredictable fluctuations in the environment. In general, however, both types of error can contribute to a cue's unreliability. As discussed in Section 4.2.2, many kinds of developmental cues can be simply described as a composition of these two types of error. Error at the population level leads to variation in the distribution of cues, as described by the population-level cue q ; individuals then draw cues c independently from that distribution.

What does the optimal developmental strategy for a cue with both kinds of error look like? As described in Section 4.2.3, the long-term growth rate of any developmental strategy depends on the phenotypic diversity it generates, within and between generations. This is described via the matrix \mathbf{CG} , which captures the effective population-level strategy, $\bar{g}(x|q)$. What if individuals could directly observe the population-level cue q ? Let \mathbf{G}_q^* indicate the optimal developmental strategy in response to the population-level component of the cue, q . From Section 4.3.1 we know that this strategy tries to match average specialization levels $\mathbf{G}_q^* \mathbf{S}$ to the conditional probabilities of environments $\tilde{\mathbf{Q}}$. We would like to find the strategy \mathbf{G}^* which, by responding simply to individually received developmental cues c , can achieve the optimal population-level strategy:

$$\begin{array}{l} \text{optimal strategy,} \\ \text{general cue} \end{array} \quad \mathbf{CG}^* \mathbf{S} = \tilde{\mathbf{Q}}. \quad (4.13)$$

Whether this is possible depends on whether the conditional probability of environments $\tilde{\mathbf{Q}}$ falls within the range jointly defined by the specialization matrix \mathbf{S} and the individual errors \mathbf{C} . For example, if $\Pr(c_1|q_1)s(e_1|x_1) + \Pr(c_2|q_1)s(e_1|x_2) = \Pr(e_1|q_1)$, and $\Pr(c_1|q_2)s(e_2|x_1) + \Pr(c_2|q_2)s(e_2|x_2) = \Pr(e_2|q_2)$ then a conditional strategy that responds to c_1 with x_1 and to c_2 with x_2 would create just

the right levels of specialization. This is illustrated in Figure 4.6, for $\mathbf{C} = \begin{pmatrix} 0.75 & 0.25 \\ 0.15 & 0.85 \end{pmatrix}$ and the usual $\mathbf{S} = \begin{pmatrix} 0.70 & 0.3 \\ 0.20 & 0.8 \end{pmatrix}$ (point shown in white.) The two pure strategies and the two conditional strategies together define the entire range of strategies that can be used. The average specialization in each environment $\bar{s}(e|q) = \sum_c \Pr(c|q) \sum_x g(x|c) s(e|x)$ achieved for each of those strategies define the four points of a parallelogram in which a fully stochastic developmental strategy is optimal (central region, shown in orange.)

The optimal growth rate for a cue with both levels of error can now be found, at least for the cases where the optimal developmental strategy is fully stochastic. If Equation 4.13 can be satisfied, the optimal growth rate is obtained by substituting into Equation 4.5:

$$\begin{array}{l} \text{optimal} \\ \text{growth rate,} \\ \text{general cue} \end{array} \quad r_B(\mathbf{G}^*) = \mathbf{Q} : \log \mathbf{C}\mathbf{G}^*\mathbf{S}\mathbf{D} = \mathbf{Q} : \log \tilde{\mathbf{Q}}\mathbf{D}. \quad (4.14)$$

If the optimal specialization levels cannot be attained with a fully stochastic strategy, then the best partially stochastic strategy will be used. The best partially stochastic strategy is the one that minimizes the Kullback-Leibler divergence between the optimal specialization profile $\Pr(e|q)$, as given by $\tilde{\mathbf{Q}}$, and the achieved one $\bar{s}(e|q)$, as given by $\mathbf{C}\mathbf{G}\mathbf{S}$.

When the optimal specialization profile is attainable, the growth rate it produces is exactly the same one that could be achieved if all individuals received the population-level cue q . The value is therefore precisely the Shannon information contained by that population-level cue about the environment:

$$\begin{array}{l} \text{adaptive value,} \\ \text{general cue} \end{array} \quad V_B(C) = V_P(Q) = I(E; Q). \quad (4.15)$$

This is illustrated in Figure 4.7, where the value of a cue with error at both levels is compared to the Shannon information in the population-level cue (center region in black, lower right; compare to Figure 4.3.) The population-level information in the cue is an upper limit on the value. On the other hand, because the information in the individual cue $I(E; C)$ is always less than the information in the population level cue $I(E; Q)$, a cue may be even more valuable than the information it gives the individual (Figure 4.7, upper right.)

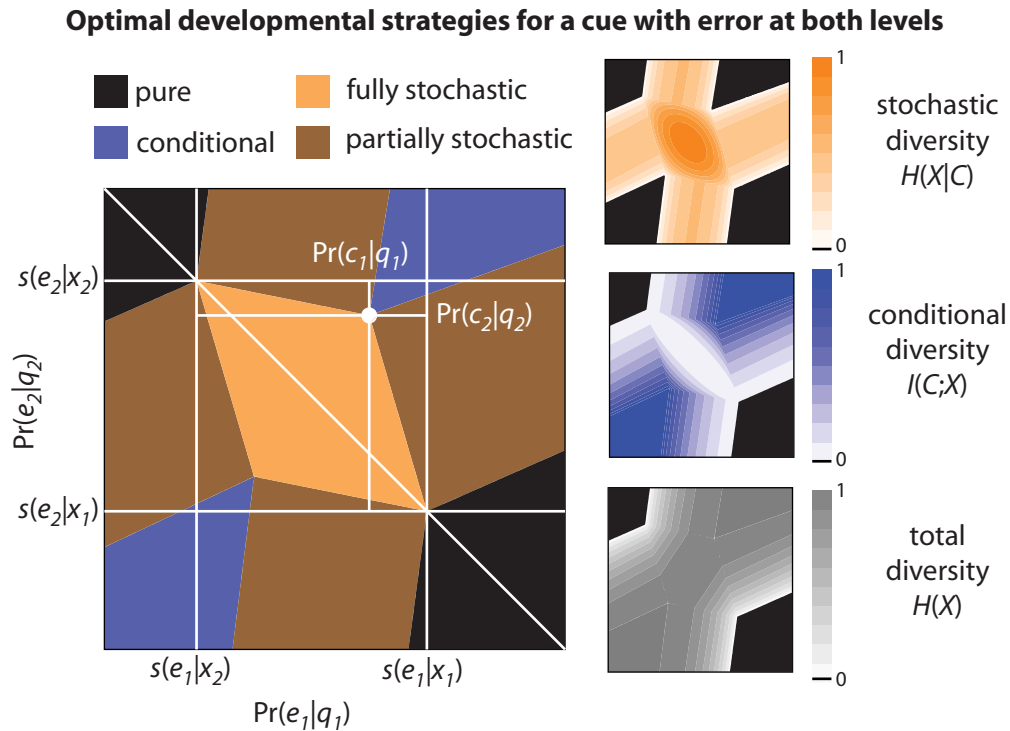


Figure 4.6: Optimal developmental strategies in response to a cue with varying amounts of error at the population level, given a fixed amount of individual-level error. For comparison to the examples for population-level error alone (Figure 4.2) and individual-level error alone (Figure 4.4), the same specialization matrix \mathbf{S} is used, with $s(e_1|x_1) = 0.7$ and $s(e_2|x_2) = 0.8$. However, now the individual-level error is fixed at $Pr(c_1|q_1) = 0.75$ and $Pr(c_2|q_2) = 0.85$, while the population-level error $Pr(e|q)$ varies. As in Figure 4.2, we categorize the optimal developmental strategies according to type in the left panel. As for the case with population-level error alone illustrated in Figure 4.2, the specialization matrix \mathbf{S} limits the range of conditional environmental probabilities in which a fully stochastic strategy can be used to achieve the optimal specialization profile. However, the addition of individual-level error, as represented by the matrix \mathbf{C} , further constrains that range to a parallelogram in the center, shown in orange. In the column at right, the optimal strategy is characterized according to how much phenotypic diversity it produces, through stochastic (top) and conditional (middle) regulatory mechanisms and in total (bottom.) As in Figure 4.1, more intense colors indicate greater phenotypic diversity. Notice the large regions of zero diversity, shown in black.

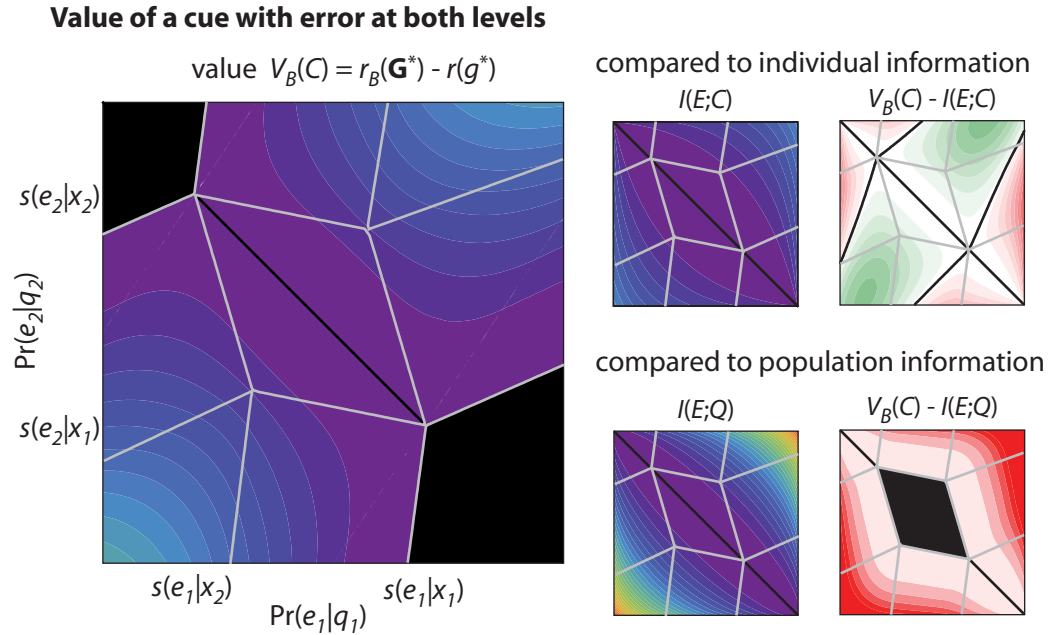


Figure 4.7: The value of a cue with varying amounts of error at the population level, given a fixed amount of individual-level error. The value of the cue, $V_B(C)$, is equal to the difference in the optimal growth rates of two lineages: one that uses the cue with error at both levels, and one that uses no cue at all. To illustrate the relationship of the amount of information a cue contains about the environment to the value it confers, we return to the example of Figure 4.6, where $s(e_1|x_1) = 0.7$ and $s(e_2|x_2) = 0.8$. As in Figure 4.6, the individual-level error is fixed at $\Pr(c_1|q_1) = 0.75$ and $\Pr(c_2|q_2) = 0.85$, while the population-level error $\Pr(e|q)$ varies. At left is a contour plot of the value of the cue, as the population-level reliability varies. For comparison, the amount of information in an individual cue, as measured by the Shannon information $I(E;C)$, is plotted on the upper right, using the same color scheme as the value; next to it is the difference between that and the value of the cue. The amount of information in the population-level cue $I(E;Q)$ is plotted on the lower right; next to it is the difference between that and the value of the cue. While the value of the cue can sometimes be greater than the information the individual gains about the environment, $I(E;C)$ (upper right), the value can never exceed the amount of information gained by the entire population about the environment, $I(E;Q)$ (lower right.) In fact, for a large region—wherever bet-hedging is still necessary—the value is exactly equal to the amount of population-level information (black region, lower right.) As usual, the regions where different types of strategies are optimal, as calculated in Figure 4.6, are outlined in gray.

4.4 Discussion

Phenotypic plasticity plays an important role in adaptation to variable environments. Most of the work that has been done on this topic has focused on adaptation to individual-level risk. For example, many species show induced morphological defenses in response to the presence of predators (e.g. barnacles, Lively 1986b; freshwater snails, DeWitt 1998; tadpoles, McCollum and Buskirk 1996.) Theoretical models indicate that in such cases a completely conditional response can evolve even in response to imperfect cues (Lively, 1986a; Sultan and Spencer, 2002). Phenotypic plasticity as a response to population-level risk, however, shows very different patterns. The pioneering model of Cohen (1967), based on patterns of seed dormancy in desert annual plants, showed that cues which help predict germination success should simply change the fraction of seeds which remain dormant. Empirical work on the germination response of such plants has shown that the germination fraction indeed varies with several predictive factors like the amount of rain and the temperature (Philippi, 1993; Clauss and Venable, 2000; Adondakis and Venable, 2004). More general theoretical models show that if population-level risk selects for random phenotypic diversity, predictive cues will modify the optimal amount of diversity (Haccou and Iwasa, 1995; DeWitt and Langerhans, 2004).

Population-level risk creates a need for phenotypic diversity, but random production of phenotypes can limit the growth rate of a lineage. The cost of having to bet-hedge—low fitness for those individuals whose phenotype does not match their environment—is directly related to the amount of uncertainty about the environment. The value in a cue about population-level risk stems from its ability to reduce environmental uncertainty, and thus reduce the cost of bet-hedging (Bergstrom and Lachmann, 2004). The Shannon information in a cue about the environment expresses how much the environmental uncertainty is reduced, once the cue is known. This explains the connection between the increased growth rate made possible by using a cue, and the amount of information in that cue. A particularly compelling example of the link between growth rate and information has been shown for the case of bacterial persistence: *E. coli* can switch between a slow-growing, antibiotic-resistance phenotype, and a fast-growing non-resistant type. The switching rate (a form of phenotypic plasticity) seems to be adapted to the frequency with which antibiotics are encountered; furthermore, it may be altered if the bacteria can sense the presence of antibiotics (Thattai and van

Oudenaarden, 2004; Kussell et al., 2005; Kussell and Leibler, 2005).

Given this tight link between the information in a developmental cue about population-level risk and its adaptive value, how is it that, as shown here, a developmental cue can sometimes increase the growth rate even past this limit? The explanation lies in the way the cue is distributed between individuals within the same generation. An individual's uncertainty about the environment it will encounter is reduced by the developmental cue it receives. When all individuals receive the same cue, that reduction in uncertainty is completely at the population level. The adaptive value of the cue is thus precisely the information it contains. However, when different individuals receive different cues, the population-level uncertainty about the environment and the individual-level uncertainty about the environment are not equivalent. How much uncertainty remains about the environment at the population level—and thus serves to drive bet-hedging—depends on how much information can be obtained by considering the entire population's developmental cues together. The cost of incomplete information imposed by having to bet-hedge is thus dependent on the population's uncertainty about the environment, which is generally less than the individual's.

This observation that the value of a cue depends on the information available to the entire population seems to suggest that individual-level error in a cue does not matter at all. One might imagine that if all individuals in the population could observe each other's developmental cues, thus improving their ability to predict environmental change, the long-term growth rate of the lineage could be increased. However, under certain conditions, such information sharing would, surprisingly, have no benefit at all. As long as the population-level uncertainty is large enough, individuals can use all of the individual-level uncertainty in the cue they receive as an external source of adaptive stochastic variation. When this is true, the individual-level error has absolutely no cost, because without it the optimal strategy would still create just as much phenotypic diversity—with internally generated developmental noise.

Although the importance of the distinction between population-level risk and individual-level risk is well recognized where fitness functions are concerned (e.g. Seger and Brockmann, 1987; Frank and Slatkin, 1990; Moran, 1992; Robson, 1996), little attention has been paid to this distinction in the distribution of cues about risk. In general, models of plasticity as an adaptation to individual-level risk posit individual-level cues (Lively, 1986a; Moran, 1992), while models of plasticity as an adaptation to population-level risk consider only population-level cues (Cohen, 1967;

Moran, 1992; Haccou and Iwasa, 1995). The focus is on how individuals use the information in the cue to create correlations between phenotype and environment. An exception to this is the idea that one mechanism for generating adaptive random phenotypic diversity is to respond conditionally to unpredictable features of the developmental environment (Cooper and Kaplan, 1982; Simons and Johnston, 2006). A cue that has no information at all could still be used as a coin-flip to generate random phenotypic diversity, as long as its error is completely at the individual level. We show that it may often be more fruitful to think of cues about population-level risk as combining elements of both these approaches. To the extent that the cue reduces the population-level uncertainty about the environment, it can increase the growth rate by reducing the need for bet-hedging. At the same time, if bet-hedging is still adaptive, any individual-level error in the cue may be exploited as a source of phenotypic diversity.

Appendices

A Levels of error in a developmental cue

Using a powerful theorem from probability theory called de Finetti's theorem, many kinds of correlation structures in developmental cues can be thought of as stemming entirely from error at the individual and population levels. If each individual i in the population receives a developmental cue c_i , we are interested in the way that the joint distribution of developmental cues over all individuals in the population depends on the type of environment e , that is, $\Pr(c_1, c_2, \dots, c_n | e)$. We would like to rewrite that joint distribution as follows: $\Pr(c_1, c_2, \dots, c_n | e) = \sum_q \Pr(q | e) \prod_i \Pr(c_i | q)$, where $\Pr(q | e)$ describes the population-level error in the cue, and $\Pr(c_i | q)$ describes the individual-level error in the cue. The variable q —the population-level cue—represents the type of generation, in terms of how the cues depend on the environment. It is not obvious that this kind of breakdown should always be possible; in particular, why should the individual cues be conditionally independent, given the type of generation? According to de Finetti's theorem, only one assumption is necessary: the cues that different individuals receive in the same generation, c_1, \dots, c_n , are what is called *exchangeable variables*.

Exchangeability is a property of identically distributed random variables which is weaker than independence, but is nonetheless strong enough to be useful. A finite sequence of random variables

A_1, \dots, A_n is exchangeable if every permutation of these variables has the same probability distribution, and an infinite sequence is exchangeable if every finite subsequence is exchangeable. De Finetti's theorem states that any infinite, exchangeable sequence of binary random variables can be understood as a draw from some probability distribution of a single random parameter Θ , followed by a sequence of independent Bernoulli trials weighted according to the parameter θ . That is,

$$p(a_1, \dots, a_n) = \int_{\theta} p(\theta) \theta^k (1 - \theta)^{n-k} d\theta$$

where k is the number of ones in the sequence a_1, \dots, a_n (see e.g. Feller, 1966).

One way to picture this process is to imagine a box of coins with different biases; we choose a coin from this box, and then flip that same coin over and over again. In this case the “population-level” uncertainty given by $p(\theta)$ is the uncertainty about which coin was chosen, while the “individual-level” uncertainty given by $p(a|\theta)$ is the uncertainty about individual coin-flips, given which coin was chosen. The remarkable implication of the theorem is that a rather general class of sequences of random variables can be viewed in exactly this light. Furthermore, the theorem can be easily generalized to discrete-valued (rather than only binary) random variables (Hewitt and Savage, 1955). The theorem guarantees that generations can be classified into different types, according to a parameter q which describes the probability distribution of individual cues. Conditional on the type of generation, the individual developmental cues are independent and identically distributed.

BIBLIOGRAPHY

- ADONDAKIS, S. AND VENABLE, D. 2004. Dormancy and germination in a guild of Sonoran Desert annuals. *Ecology* 85:2582–2590.
- ALLEN, C. AND SAIDEL, E. 1998. The evolution of reference, pp. 183–203. In D. D. Cummins and C. Allen (eds.), *The Evolution of Mind*. Oxford university Press, New York.
- AVERY, S. V. 2006. Microbial cell individuality and the underlying sources of heterogeneity. *Nature Reviews Microbiology* 4:577–587.
- BALKENBORG, D. 1994. Strictness and evolutionary stability. Center for Rationality and Interactive Decision Theory, Jerusalem, Discussion Paper 52.
- BALKENBORG, D. AND SCHLAG, K. H. 2001. Evolutionarily stable sets. *Int J Game Theory* 29:571–595.
- BERG, H. AND REDBO-TORSTENSSON, P. 1998. Cleistogamy as a bet-hedging strategy in *Oxalis acetosella*, a perennial herb. *Journal of Ecology* 86:491–500.
- BERGSTROM, C. T. AND GODFREY SMITH, P. 1998. On the evolution of behavioral heterogeneity in individuals and populations. *Biology and Philosophy* 13:205–231.
- BERGSTROM, C. T. AND LACHMANN, M. 2004. Shannon information and biological fitness. In *IEEE Information Theory Workshop 2004*, pp. 50–54. IEEE. (See also arXiv.org:q-bio/0510007).
- BICKERTON, D. 1990. *Language and Species*. University of Chicago Press, Chicago.
- BOYCE, M. S. AND PERRINS, C. M. 1987. Optimizing great tit clutch size in a fluctuating environment. *Ecology* 68:142–153.
- BRADFORD, M. J. AND ROFF, D. A. 1997. An empirical model of diapause strategies of the cricket *Allonemobius socius*. *Ecology* 78:442–451.
- BULL, J. J. 1987. Evolution of phenotypic variance. *Evolution* 41:303–315.
- CAPINERA, J. L. 1979. Qualitative variation in plants and insects: Effect of propagule size on ecological plasticity. *American Naturalist* 114:350–361.
- CHENEY, D. L. AND SEYFARTH, R. M. 1990. *How monkeys see the world*. University of Chicago Press, Chicago.
- CLAUSS, M. J. AND VENABLE, D. L. 2000. Seed germination in desert annuals: An empirical test of adaptive bet-hedging. *American Naturalist* 155:168–186.
- COHEN, D. 1966. Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology* 12:119–129.

- COHEN, D. 1967. Optimizing reproduction in a randomly varying environment when a correlation may exist between the conditions at the time a choice has to be made and the subsequent outcome. *Journal of Theoretical Biology* 16:1–14.
- COOPER, W. S. AND KAPLAN, R. H. 1982. Adaptive “coin-flipping”: a decision-theoretic examination of natural selection for random individual variation. *Journal of Theoretical Biology* 94:135–151.
- COVER, T. M. AND THOMAS, J. A. 1991. Elements of Information Theory. Wiley Series in Telecommunications. John Wiley & Sons, New York.
- CRUMP, M. L. 1981. Variation in propagule size as a function of environmental uncertainty for tree frogs. *American Naturalist* 117:724–737.
- DANFORTH, B. N. 1999. Emergence dynamics and bet hedging in a desert bee, *Perdita portalis*. *Proceedings of the Royal Society of London, Series B* 266:1985–1994.
- DE MAZANCOURT, C. AND DIECKMANN, U. 2004. Trade-off geometries and frequency-dependent selection. *American Naturalist* 164:765–778.
- DEMPSTER, E. R. 1955. Maintenance of genetic heterogeneity. *Cold Spring Harbor Symposia on Quantitative Biology* 20:25–32.
- DENVER, R. J., MIRHADI, N., AND PHILLIPS, M. 1998. Adaptive plasticity in amphibian metamorphosis: Response of *Scaphiopus hammonii* tadpoles to habitat desiccation. *Ecology* 79:1859–1872.
- DEWITT, T. J. 1998. Costs and limits of phenotypic plasticity: Tests with predator-induced morphology and life history in a freshwater snail. *Journal of Evolutionary Biology* 11:465–480.
- DEWITT, T. J. AND LANGERHANS, R. B. 2004. Integrated solutions to environmental heterogeneity: Theory of multimoment reaction norms, pp. 98–111. In T. J. DeWitt and S. M. Scheiner (eds.), Phenotypic Plasticity: Functional and Conceptual Approaches. Oxford University Press, New York.
- DEWITT, T. J., ROBINSON, B. W., AND WILSON, D. S. 2000. Functional diversity among predators of a freshwater snail imposes an adaptive trade-off for shell morphology. *Evolutionary Ecology Research* 2:129–148.
- DEWITT, T. J., SIH, A., AND WILSON, D. S. 1998. Costs and limits of phenotypic plasticity. *Trends in Ecology and Evolution* 13:77–81.
- DITTUS, W. P. 1984. Toque macaque food calls: semantic communication concerning food distribution in the environment. *Anim Behav* 32:470–477.
- DONALDSON, M. C., LACHMANN, M., AND BERGSTROM, C. T. 2007. The evolution of functionally referential meaning in a structured world. *Journal of Theoretical Biology* 246:225–233.

- DONALDSON-MATASCI, M. C., LACHMANN, M., AND BERGSTROM, C. T. 2008. Phenotypic diversity as an adaptation to environmental uncertainty. *Evolutionary Ecology Research* 10:493–515.
- EINUM, S. AND FLEMING, I. A. 2004. Environmental unpredictability and offspring size: conservative versus diversified bet-hedging. *Evolutionary Ecology Research* 6:443–455.
- ESHEL, I. AND FELDMAN, M. W. 1984. Initial increase of new mutants and some continuity properties of ESS in two-locus systems. *Am Nat* 124:631–640.
- EVANS, C. S. AND EVANS, L. 1999. Chicken food calls are functionally referential. *Anim Behav* 58:307–319.
- EVANS, C. S., EVANS, L., AND MARLER, P. 1993. On the meaning of alarm calls: functional reference in an avian vocal system. *Anim Behav* 46:23–38.
- EVANS, M. E. K., FERRIÈRE, R., KANE, M. J., AND VENABLE, D. L. 2007. Bet hedging via seed banking in desert evening primroses (*Oenothera*, Onagraceae): Demographic evidence from natural populations. *American Naturalist* 169:184–194.
- FELLER, W. 1966. An Introduction to Probability Theory and Its Applications, volume II. Wiley, New York.
- FELSENSTEIN, J. 1976. The theoretical population genetics of variable selection and migration. *Annual Review of Genetics* 10:253–280.
- FRANK, S. A. AND SLATKIN, M. 1990. Evolution in a variable environment. *American Naturalist* 136:244–260.
- FRANKINO, W. A. AND PFENNIG, D. W. 2001. Condition-dependent expression of trophic polyphenism: effects of individual size and competitive ability. *Evolutionary Ecology Research* 3:939–951.
- GABRIEL, W., LUTTBEG, B., SIH, A., AND TOLLRIAN, R. 2005. Environmental tolerance, heterogeneity, and the evolution of reversible plastic responses. *American Naturalist* 166:339–353.
- GILLESPIE, J. 1973. Polymorphism in random environments. *Theoretical Population Biology* 4:193–195.
- GODFREY-SMITH, P. 1996. Complexity and the Function of Mind in Nature. Cambridge Studies in Philosophy and Biology. Cambridge University Press, Cambridge.
- GOODMAN, D. 1984. Risk spreading as an adaptive strategy in iteroparous life histories. *Theoretical Population Biology* 25:1–20.
- GOUZOULES, S., GOUZOULES, H., AND MARLER, P. 1984. Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Anim Behav* 32:182–193.
- GRAFEN, A. 1999. Formal darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proceedings of the Royal Society of London, Series B* 266:799–803.

- GRANT, P. R. AND GRANT, B. R. 1997. Genetics and the origin of bird species. *Proc Natl Acad Sci USA* 94:7768–7775.
- GRICE, H. P. 1957. Meaning. *Philos Rev* 66:377–388.
- HACCOU, P. AND IWASA, Y. 1995. Optimal mixed strategies in stochastic environments. *Theoretical Population Biology* 47:212–243.
- HALDANE, J. AND JAYAKAR, S. 1962. Polymorphism due to selection of varying direction. *Journal of Genetics* 58:237–242.
- HALKETT, F., HARRINGTON, R., HULLÉ, M., KINDLMANN, P., MENU, F., RISPE, C., AND PLANTEGENEST, M. 2004. Dynamics of production of sexual forms in aphids: Theoretical and experimental evidence for adaptive “coin-flipping” plasticity. *American Naturalist* 163:E112–E125.
- HAMMERSTEIN, P. 1996. Darwinian adaptation, population genetics and the streetcar theory of evolution. *J Math Biol* 34:511–532.
- HASSALL, M., WALTERS, R. J., TELFER, M., AND HASSALL, M. R. J. 2006. Why does a grasshopper have fewer, larger offspring at its range limits? *Journal of Evolutionary Biology* 19:267–276.
- HAUSER, M. D. 1998. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Anim Behav* 55:1647–1658.
- HAUSER, M. D., CHOMSKY, N., AND FITCH, W. T. 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science* 298:1569–1579.
- HEWITT, E. AND SAVAGE, L. J. 1955. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society* 80:470–501.
- HOPPER, K. R. 1999. Risk-spreading and bet-hedging in insect population biology. *Annual Review of Entomology* 44:535–560.
- HURFORD, J. R. 1989. Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua* 77:187–222.
- KAPLAN, R. H. AND COOPER, W. S. 1984. The evolution of developmental plasticity in reproductive characteristics: An application of the “adaptive coin-flipping” principle. *American Naturalist* 123:393–410.
- KELLY, J. L. 1956. A new interpretation of information rate. *Bell System Technical Journal* 35:917–926.
- KOOPS, M. A., HUTCHINGS, J. A., AND ADAMS, B. K. 2003. Environmental predictability and the cost of imperfect information: influences on offspring size variability. *Evolutionary Ecology Research* 5:29–42.

- KUSSELL, E., KISHONY, R., BALABAN, N., AND LEIBLER, S. 2005. Bacterial persistence: A model of survival in changing environments. *Genetics* 169:1807–1814.
- KUSSELL, E. AND LEIBLER, S. 2005. Phenotypic diversity, population growth and information in fluctuating environments. *Science* 309:2075–2078.
- LACHMANN, M. AND BERGSTROM, C. T. 1998. Signalling among relatives. II. Beyond the tower of Babel. *Theor Pop Biol* 54:146–160.
- LACHMANN, M., SZÁMADÓ, S., AND BERGSTROM, C. T. 2001. Cost and conflict in animal signals and human language. *Proc Nat Acad Sci USA* 98:13189–13194.
- LANE, S. J. AND MAHONY, M. J. 2002. Larval anurans with synchronous and asynchronous development periods: Contrasting responses to water reduction and predator presence. *Journal of Animal Ecology* 71:780–792.
- LEIMAR, O. 2005. The evolution of phenotypic polymorphism: Randomized strategies versus evolutionary branching. *American Naturalist* 165:669–681.
- LEVENE, H. 1953. Genetic equilibrium when more than one ecological niche is available. *The American Naturalist* 87:331–333.
- LEVINS, R. 1962. The theory of fitness in a heterogeneous environment. I. The fitness set and adaptive function. *American Naturalist* 96:361–373.
- LEVINS, R. 1964. Theory of fitness in a heterogeneous environment. III. The response to selection. *Journal of Theoretical Biology* 7:224–240.
- LEVINS, R. 1965. The theory of fitness in a heterogeneous environment. V. Optimal genetic systems. *Genetics* 52:891–904.
- LEVINS, R. 1968. Evolution in changing environments. Princeton University Press, Princeton, New Jersey.
- LEWIS, D. K. 1969. Convention: A Philosophical Study. Harvard University Press, Cambridge, MA.
- LIVELY, C. M. 1986a. Canalization versus developmental conversion in a spatially-variable environment. *American Naturalist* 128:561–572.
- LIVELY, C. M. 1986b. Predator-induced shell dimorphism in the acorn barnacle *Chthamalus anisopoma*. *Evolution* 40:232–242.
- MACEDONIA, J. M. 1990. What is communicated in the antipredator calls of lemurs: evidence from playback experiments with ringtailed and ruffed lemurs. *Ethology* 86:177–190.
- MACEDONIA, J. M. AND EVANS, C. S. 1993. Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* 93:177–197.

- MANSER, M. B. 2001. The acoustic structure of suricate alarm calls varies with predator type and the level of response urgency. *Proc Roy Soc Lond B* 268:2315–2324.
- MARLER, P., EVANS, C. S., AND HAUSER, M. D. 1992. Animal signals: motivational, referential, or both?, pp. 66–86. In H. Papoušek, U. Jürgens, and M. Papoušek (eds.), *Nonverbal vocal communication: comparative and developmental approaches*. Cambridge University Press, Cambridge.
- MAYNARD SMITH, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- MAYNARD SMITH, J. AND HARPER, D. 2003. *Animal Signals*. Oxford University Press, New York.
- MAYNARD SMITH, J. AND PRICE, G. R. 1973. The logic of animal conflict. *Nature* 246:15–18.
- MCCOLLUM, S. A. AND BUSKIRK, J. V. 1996. Costs and benefits of a predator-induced polyphenism in the gray treefrog *Hyla chrysoscelis*. *Evolution* 50:583–593.
- MCNAMARA, J. M. 1995. Implicit frequency dependence and kin selection in fluctuating environments. *Evolutionary Ecology* 9:185–203.
- MENU, F., ROEBUCK, J.-P., AND VIALA, M. 2000. Bet-hedging diapause strategies in stochastic environments. *American Naturalist* 155:724–734.
- METZ, J. A. J., NISBET, R. M., AND GERITZ, S. A. H. 1992. How should we define ‘fitness’ for general ecological scenarios? *Trends in Ecology and Evolution* 7:198–202.
- MORAN, N. A. 1992. The evolutionary maintenance of alternative phenotypes. *American Naturalist* 139:971–989.
- MOREY, S. R. AND REZNICK, D. N. 2004. The relationship between habitat permanence and larval development in California spadefoot toads: Field and laboratory comparisons of developmental plasticity. *Oikos* 104:172–190.
- MOYNIHAN, M. 1970. Control, suppression, decay, disappearance and replacement of displays. *J Theor Biol* 29:85–112.
- MURPHY, G. I. 1968. Pattern in life history and the environment. *American Naturalist* 102:391–403.
- NEWMAN, R. A. 1992. Adaptive plasticity in amphibian metamorphosis. *BioScience* 42:671–678.
- NOWAK, M. A. AND KRAKAUER, D. C. 1999. The evolution of language. *Proc Nat Acad Sci USA* 96:8028–8023.
- NOWAK, M. A., PLOTKIN, J. B., AND DRESS, A. 1999. An error limit for the evolution of language. *Proc R Soc Lond B Biol Sci* 266:2131–2136.

- OWREN, M. J. AND RENDALL, D. 2001. Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evol Anthropology* 10:58–71.
- PADILLA, D. K. AND ADOLPH, S. C. 1996. Plastic inducible morphologies are not always adaptive: The importance of time delays in a stochastic environment. *Evolutionary Ecology* 10:105–117.
- PHILIPPI, T. 1993. Bet-hedging germination of desert annuals: Variation among populations and maternal effects in *Lepidium lasiocarpum*. *American Naturalist* 142:488–507.
- PHILIPPI, T. AND SEGER, J. 1989. Hedging one's evolutionary bets, revisited. *Trends in Ecology and Evolution* 4:41–44.
- PHILIPPI, T., SIMOVICH, M., BAUDER, E., AND MOORAD, J. 2001. Habitat ephemerality and hatching fractions of a diapausing anostracan (Crustacea : Branchiopoda). *Israel Journal of Zoology* 47:387–395.
- PINKER, S. AND JACKENDOFF, R. 2005. The faculty of language: what's special about it? *Cognition* 95:201–236.
- QUINE, W. V. 1960. *Word and Object*. MIT Press, Cambridge, MA.
- RELYEA, R. A. 2002. Costs of phenotypic plasticity. *American Naturalist* 159:272–282.
- RELYEA, R. A. 2003. Predators come and predators go: The reversibility of predator-induced traits. *Ecology* 84:1840–1848.
- ROBINSON, S. R. 1981. Alarm communication in Belding's ground squirrels. *Z Tierpsychol* 56:150–168.
- ROBSON, A. J. 1996. A biological basis for expected utility. *Journal of Economic Theory* 68:397–424.
- RUEFFLER, C., VAN DOOREN, T. J. M., AND METZ, J. A. J. 2004. Adaptive walks on changing landscapes: Levins' approach extended. *Theoretical Population Biology* 65:165–178.
- SAIAH, H. AND PERRIN, N. 1990. Autumnal vs. spring hatching in the fairy shrimp *Siphonophanes grubii* (Dybowski) (Crustacea, Anostraca): diversified bet-hedging strategy? *Functional Ecology* 4:769–775.
- SASAKI, A. AND ELLNER, S. 1995. The evolutionarily stable phenotype distribution in a random environment. *Evolution* 49:337–350.
- SEGER, J. AND BROCKMANN, H. J. 1987. What is bet-hedging?, pp. 182–211. In P. Harvey and L. Partridge (eds.), *Oxford Surveys in Evolutionary Biology*, volume 4. Oxford University Press, Oxford.
- SELTEN, R. 1980. A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology* 84:93–101.

- SEYFARTH, R. M., CHENEY, D. L., AND MARLER, P. 1980. Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim Behav* 28:1070–1094.
- SIMONS, A. M. AND JOHNSTON, M. O. 2003. Suboptimal timing of reproduction in *Lobelia inflata* may be a conservative bet-hedging strategy. *Journal of Evolutionary Biology* 16:233–243.
- SIMONS, A. M. AND JOHNSTON, M. O. 2006. Environmental and genetic sources of diversification in the timing of seed germination: Implications for the evolution of bet hedging. *Evolution* 60:2280–2292.
- SKÚLASON, S. AND SMITH, T. B. 1995. Resource polymorphisms in vertebrates. *Trends in Ecology and Evolution* 10:366–370.
- SLATKIN, M. AND LANDE, R. 1976. Niche width in a fluctuating environment—density independent model. *American Naturalist* 110:31–55.
- SMITH, T. B. 1993. Disruptive selection and the genetic basis of bill size polymorphism in the African finch *Pyrenestes*. *Nature* 363:618–620.
- SMITS, W. K., KUIPERS, O. P., AND VEENING, J.-W. 2006. Phenotypic variation in bacteria: the role of feedback regulation. *Nature Reviews Microbiology* 4:259–271.
- STEARNS, S. C. 1989. The evolutionary significance of phenotypic plasticity—phenotypic sources of variation among organisms can be described by developmental switches and reaction norms. *BioScience* 39:436–445.
- SULTAN, S. AND SPENCER, H. 2002. Metapopulation structure favors plasticity over local adaptation. *American Naturalist* 160:271–283.
- TEBBICH, S., TABORSKY, M., FESSL, B., DVORAK, M., AND WINKLER, H. 2004. Feeding behavior of four arboreal Darwin's finches: Adaptations to spatial and seasonal variability. *Condor* 106:95–105.
- TEMPLETON, A. R. AND ROTHMAN, E. D. 1981. Evolution in fine-grained environments. II. Habitat selection as a homeostatic mechanism. *Theoretical Population Biology* 19:326–340.
- THATTAI, M. AND VAN OUDENAARDEN, A. 2004. Stochastic gene expression in fluctuating environments. *Genetics* 167:523–530.
- THOMAS, B. 1985. On evolutionarily stable sets. *J Math Biol* 22:105–115.
- TRAPA, P. E. AND NOWAK, M. A. 2000. Nash equilibria for an evolutionary language game. *J Math Biol* 41:172–188.
- TULJAPURKAR, S. 1990. Population Dynamics in Variable Environments. Lecture Notes in Biomathematics. Springer-Verlag, Berlin.
- VAN BUSKIRK, J. 2002. Phenotypic lability and the evolution of predator-induced plasticity in tadpoles. *Evolution* 56:361–370.

- VAN KLEUNEN, M., FISCHER, M., AND SCHMID, B. 2000. Costs of plasticity in foraging characteristics of the clonal plant *ranunculus reptans*. *Evolution* 54:1947–1955.
- VENABLE, D. L. 2007. Bet hedging in a guild of desert annuals. *Ecology* 88:1086–1090.
- VIA, S. AND LANDE, R. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39:505–522.
- WAAS, J. R. 1991. The risks and benefits of signaling aggressive motivation — a study of cave-dwelling little blue penguins. *Behav Ecol Sociobiol* 29:139–146.
- WALTERS, J. R. 1990. Anti-predatory behavior of lapwings: field evidence of discriminative abilities. *Wilson Bull* 102:49–70.
- WÄRNERYD, K. 1993. Cheap talk, coordination, and evolutionary stability. *Games Econom Behav* 5:532–546.
- WERNER, T. AND SHERRY, T. 1987. Behavioral feeding specialization in *Pinaroloxias inornata*, the “Darwin’s Finch” of Cocos Island, Costa Rica. *Proceedings of the National Academy of Sciences USA* 84:5506–5510.
- WEST-EBERHARD, M. J. 1989. Phenotypic plasticity and the origins of diversity. *Annual Reviews in Ecology and Systematics* 20:249–278.
- WEST-EBERHARD, M. J. 2003. Developmental plasticity and evolution. Oxford University Press, Oxford.
- WILSON, D. S. AND YOSHIMURA, J. 1994. On the coexistence of specialists and generalists. *American Naturalist* 144:692–707.
- YOSHIMURA, J. AND CLARK, C. W. 1991. Individual adaptations in stochastic environments. *Evolutionary Ecology* 5:173–192.
- YOSHIMURA, J. AND JANSEN, V. A. A. 1996. Evolution and population dynamics in stochastic environments. *Researches in Population Ecology* 38:165–182.
- ZUBERBÜHLER, K., NOË, R., AND SEYFARTH, R. M. 1997. Diana monkey long-distance calls: messages for conspecifics and predators. *Anim Behav* 53:589–604.
- ZUIDEMA, W. 2003. Optimal communication in a noisy and heterogeneous environment. *Proceedings. Lecture Notes in Artificial Intelligence* 2801:553–563.

VITA

Matina Donaldson-Matasci earned a B.A. in Mathematics from Reed College in 1996. She then spent six years as a computer programmer for a series of dot-coms and newspapers, including the New York Times. From 2000-2002, she also attended Portland State University as a part-time student, taking several biology courses. There she discovered the field of mathematical biology, which seemed a perfect marriage of her interests and skills. She began a Ph.D. program in Biology at the University of Washington in 2002, where she was selected by the department to receive the merit-based ARCS Fellowship for incoming students. In 2004 she received a three-year Graduate Research Fellowship from the National Science Foundation. This allowed her to pursue a collaboration with researchers at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, where she spent most of her time from 2005-2008. In June 2008, she received a Ph.D. in Biology from the University of Washington.

Presentations and Publications

- Donaldson, M.C. (2003) *A game theoretic analysis of the evolution of communication systems*. Invited speaker, Workshop on the Evolution of Meaning. Institute for Logic, Language and Computation, University of Amsterdam.
- Donaldson, M.C., C.T. Bergstrom and M. Lachmann. (2004) *The evolution of referential meaning in communication systems*. Talk, Evolution Conference. Colorado State University, Fort Collins.
- Donaldson, M.C., C.T. Bergstrom and M. Lachmann. (2006) *Making the best of bad cues*. Poster, International Society for Behavioural Ecology Congress. Tours, France.
- Donaldson, M.C., M. Lachmann and C.T. Bergstrom. (2007) *The evolution of functionally referential meaning in a structured world*. Journal of Theoretical Biology 246 (2): 225-233.
- Donaldson, M.C., C.T. Bergstrom and M. Lachmann. (2007) *Two kinds of generalists: adaptation to environmental risk*. Poster, European Society for Evolutionary Biology Congress. Uppsala, Sweden.
- Donaldson-Matasci, M.C., M. Lachmann and C.T. Bergstrom. (2007) *Phenotypic diversity as an adaptation to environmental uncertainty*. Evolutionary Ecology Research 10: 493-515.