

What determines biological fitness?

The problem of the reference environment

Marshall Abrams (mabrams@uab.edu)

Department of Philosophy, University of Alabama at Birmingham

September 18, 2007

Abstract.

Organisms' environments are thought to play a fundamental role in determining their fitness and hence in natural selection. Existing intuitive conceptions of environment are sufficient for biological practice. I argue, however, that attempts to produce a general characterization of fitness and natural selection are incomplete without the help of general conceptions of what conditions are included in the environment. Thus there is a “problem of the reference environment”—more particularly, problems of specifying principles which pick out those environmental conditions which determine fitness. I distinguish various reference environment problems and propose solutions to some of them. While there has been a limited amount of work on problems concerning what I call “subenvironments”, there appears to be no earlier work on problems of what I call the “whole environment”. The first solution I propose for a whole environment problem specifies the overall environment for natural selection on a set of biological types present in a population over a specified period of time. The second specifies an environment relevant to extinction of types in a population; this kind of environment is especially relevant to certain kinds of long-term evolution.

Keywords: environment, fitness, natural selection, niche, probability

1. Introduction

Natural selection, the central concept of evolutionary theory, is usually thought to have to do with differences in fitness. Fitness, in turn, is usually thought to have to do with probabilities of survival and reproduction. Finally, it is almost always assumed that these probabilities depend on the relationship between organisms' genotypes or phenotypes and the environment. While much attention has been given to the possibility of giving a general characterization of what fitness is, little attention has been given to providing a general characterization of what an environment is. Without such a characterization, a general characterization of fitness is incomplete; by itself it may not be able to tell us which organisms are fittest, even given full information about all biologically relevant states of the world.

The purpose of this paper is to clarify this “problem of the reference environment”, to distinguish various dimensions of it, and to propose solutions to certain of its variants. Among other things, I distinguish between problems concerning “whole environments”, which include all of the conditions relevant to selection on a population, and those concerning “subenvironments”, which include more limited ranges of conditions. I propose solutions to reference environment problems concerning the whole environment; I believe that this is the first publication to do so. One of the solutions I propose (the “extinction environment”) has additional significance in that it provides a novel illustration of a way in which an “object” of both theoretical and practical importance, yet without obvious boundaries, is constituted

by its probable effects. The main illustrations of this possibility up until now were objects defined in terms of function in various senses.

After distinguishing aspects of the problem of the reference environment (§2), I'll present one kind of problem of the reference environment in more detail (§3) and propose solutions to some of its variants (§§4, 5). The concluding section (§6) includes a discussion of potential extensions and applications of my approach.

2. The problem of the reference environment

2.1. OVERVIEW

Biologists have defined “fitness” in various ways, inconsistent with each other and with various theoretical desiderata. That does not prevent them from using these definitions in useful ways in a variety of empirical and theoretical contexts. Nevertheless, it would be valuable to find a general definition of “fitness” which captured the central intuition behind its uses in a theoretically consistent way. Such an account of fitness would help to unify and make sense of biological practice and help us to understand what, in general, natural selection consists in. A number of philosophers of biology and philosophically-minded biologists have attempted to provide such general definitions of fitness (e.g. Brandon, 1990; Byerly and Michod, 1991; Rosenberg, 1994; Sober, 1984).

However, there has been little philosophical work on characterizing what conditions, in general, are included in the environment (see §2.3.1 for references).¹ This too is no impediment to biological research;

intuitions about what conditions are relevant to natural selection in particular cases are generally clear enough. A general characterization of fitness, however, is supposed to tell us what it is for some organisms to be fitter than others. That task is incomplete without a general characterization of what conditions should be included in organisms' environment; as I argue below, without a principled characterization of the environment, fitness can be evaluated relative to arbitrarily delineated sets of environmental conditions which may reverse fitness orderings. Thus a philosophical theory of fitness unsupplemented by an account of what an environment consists in will not always imply, even in principle, which organisms are fitter than others.

2.2. WHAT THE PROBLEM IS AND IS NOT

An environment is sometimes thought of as a region of space and time. That is not usually how biologists conceive of environments, and a spatiotemporal region *per se* is not the sort of thing which could determine fitness. Rather, an environment is a set of conditions, a spatiotemporal configuration of properties. Not all facts about a spatiotemporal region make a difference to fitness. For example, small variations in the stiffness of grasses might make no difference to the fitness of grazing animals. Furthermore, spatially and even temporally remote conditions such as those which generate weather patterns can make an enormous difference to an organism's fitness. Which conditions count as aspects of the environment? Since evolution takes time, and the world is not static, the conditions which determine the evolution of a population are

likely to be broader than those experienced by the population at any particular moment. That suggests a question about which of various possible conditions occurring at different times should be included in an environment.²

I call the problem of specifying what conditions determine the probabilities underlying fitness and natural selection “the problem of the reference environment”, by analogy with the problem of the reference class in philosophy of probability. More particularly, the problem of the reference environment is the problem of specifying general principles which characterize that environment which determines the fitnesses of organisms. This problem (or rather problems—see below) plausibly cuts across many different accounts of natural selection, fitness, and biological probabilities.

The problem of the reference environment is of course not a problem for processes in the world, which proceed without regard to philosophical and scientific theory; the factors which affect evolution will often be determinate whether or not we have a principled account which can distinguish them.³ Nor is the problem of the reference environment a problem about the fact that we might not know which of several alternative biological types would be likely to be successful because we lack sufficient data. The problem is that a vague conception of what would count as part of an environment gives no guidance about what facts about organisms’ surroundings *would* determine which biological types are likely to be successful. Finally, as mentioned above, the problem of the reference environment does not appear to have been a problem for research in evolutionary biology. Evolutionary biologists are usually

able to decide correctly through empirical investigation, or modeling, or more informal means what possible environmental conditions are relevant to the questions they ask. A solution to the problem of the reference environment could nevertheless codify principles which can be taken as underlying biologists' intuitions about environments. It may be that the resulting clarification would aid some aspect of biological research, but a better understanding of the foundations of evolutionary theory is sufficient motivation.

2.3. THE PROBLEMS OF THE REFERENCE ENVIRONMENT

In this section I explain that the overall problem of the reference environment can be broken into several distinct problems, varying along what I call the “grain” and “explanandum property” dimensions.

Note that in this preliminary investigation, I take reference environment problems to concern principles which specify a *set* of conditions making up the environment, ignoring important questions about how *often* combinations of conditions from this set arise; Section 6 discusses this point further.

Note that it may make sense to talk either of the fitness of an organism token or of the fitness of an organism (geno/pheno)type.⁴ My focus in this paper is on fitness of types, which is crucial to understanding natural selection conceived as change in type frequencies due to fitness differences. Some authors define type fitness in terms of token fitness (Mills and Beatty, 1979; Sober, 1984).

2.3.1. *The grain dimension*

“Environment” can refer to conditions relevant to reproductive success for all members of the population—either in the sense of conditions relevant to every member of the population, or in the wider sense of conditions relevant to at least some members of the population. I’ll use the term “whole environment” for the conditions relevant to reproductive success for at least some members of a population.

Alternatively, “environment” can refer to what I’ll call “subenvironments” (sometimes called “habitats” or “patches”). These are subsets of the whole environment relevant to reproductive success for specific subpopulations, or for members of the population during certain time periods. There are potentially many theoretically relevant ways to carve up a whole environment into subenvironments. In the limit, one can talk about particular sets of conditions relevant to the success of particular organisms—subenvironments relativized to subpopulations of size 1. Questions about how to characterize subenvironments are sometimes important. Evolution is sensitive to spatial and temporal variation (Levins, 1968; Wimsatt, 1980). Fitnesses of token organisms and of certain subpopulations may depend on particular subenvironments (Brandon, 1990, Ch. 2; Ramsey, 2006), and common models assign different fitnesses to the same types in different subenvironments (Gillespie, 1998, §3.6; Roughgarden, 1979, Ch. 13).

However, my focus in this paper will be on reference environment problems concerning the whole environment, and I’ll often drop the qualifier “whole”. Natural selection on a population has to do with effects on all members of all lineages which make up a population. It

is thus the whole environment—including any of its subenvironments—which produces natural selection. Moreover, reference environment problems concerning the whole environment are in one sense more fundamental than problems concerning subenvironments. A characterization of the whole environment limits which possible subenvironments should be considered relevant to the evolution of a given population: If the characterization implies that certain conditions are not part of the whole environment, it also implies that subenvironments including those conditions are not relevant to natural selection on the population. More correctly, a set of conditions which might have been thought to be a subenvironment for some aspect of natural selection on a given population should not be considered a subenvironment.

2.3.2. *The explanandum property dimension*

Evolution encompasses a variety of effects, including, among other things, small changes in the genotypic makeup of a population due to the disappearance of a few genes, long-term changes in the makeup of a population due to gradual accretion and disappearance of large numbers of genes affecting a single phenotypic category, and both short-term and long-term maintenance of stable genetic polymorphisms. These are effects which distributions of differentially fit traits help to explain. If fitnesses are determined in part by environmental conditions, then there are variety of evolutionary effects which environments help to explain. There may thus be different reference environment problems because distinct explananda can require distinct explanantia.⁵

I'll focus primarily on two solutions to reference environment problems for two sorts of evolutionary effects. The first solution (§4) specifies the environment relevant to the following kind of effect: Changes in frequencies of specific types all present over a specific interval of time. Here the duration of the interval over which changes occur is an essential aspect of the property whose tokening is to be explained. The second solution (§5) specifies an environment relevant to this kind of effect: Extinction of one of a specific set of related alternative types in a population (e.g. alleles at a locus). Here the duration of a time interval is not an essential aspect of what is to be explained, but the set of explananda is restricted to initial extinction events. The first concept of environment is explanatorily relevant to short-term evolution and certain aspects of long-term evolution; the second is especially relevant to questions about gradual, long-term phenotypic change involving a series of extinctions and mutations (as I'll explain below). There are undoubtedly other reference environment problems worth solving for other kinds of effects; I'll mention a few possibilities along the way, including some questions having to do with ecological niches (§6).

2.4. SOME MINIMAL ASSUMPTIONS

I'll try to remain as neutral as possible about exactly what biological fitness is, but I'll sometimes assume that fitness cannot be defined merely in terms of changes of *actual* frequencies of types in a population,⁶ and that fitness depends on probabilities whose values are affected by which conditions are included in the environment. As indicated above,

I'll assume that fitness is a property of types of individual organisms. Not everything that I say will depend on all of these assumptions; those who disagree with some of them may still find my proposals useful.

There is no one concept of population appropriate to all contexts in evolutionary theory. For example, one definition says that a population is a panmictic set of members of a species, that is, a set of organisms each of which has an equal chance of mating with any other of the opposite sex. However, this definition is not satisfied by many groups of organisms often considered to be populations, such as some groups of asexual organisms. My arguments should work with any existing concept of population according to which a population is more or less closed under the chance of recombination. That is, I am requiring that:

If there is at least a small but significant probability that there will be interbreeding and genetic recombination between an organism o or its descendants and a member of the population at some point in the interval of time under consideration, then o is also a member of the population.⁷

The reason for this requirement is that an explanation of the evolution of a population should ideally reflect those lineages which are likely to contribute genes affecting the population's evolution. Note that it's not the mere possibility of recombination at *some* future time that matters, but rather a small but significant probability of recombination within an appropriate time interval. (Spelling out which interval must wait until the end of Section 4.) Also note that this minimal requirement on populations is consistent with a population including members with no chance of recombination. The requirement may sometimes imply

that a population includes more than what is practical for empirical work or modeling, however. For example, a biologist may wish to study the evolution of a small population over a long period of time during which gene flow across the entire species is likely; in this case my conceptions of environment would be defined in terms of the entire species rather than the small population. However, practical biological work routinely involves simplifying assumptions which depart from a theoretical ideal; that is no reason not to seek a clear conception of the kinds of factors which actually cause and ideally explain and predict evolution.

3. The problem of the reference whole environment

Consider a population consisting of members of an insect species living in an arid region. The soil in this area lies on top of a layer of rock several feet below the surface. At time t_0 , each member of this population has one of two alternative phenotypes, which I will call “*deep*” and “*shallow*”. Assume that reproduction is asexual and that the *deep* and *shallow* phenotypes are determined by alternative genes at a single locus on a chromosome. The difference between the *deep* and *shallow* phenotypes is that *deep* insects tend to spend more time farther underground than do *shallow* ones, although insects of both types routinely move back and forth between the surface and the rock below. During the usual dry conditions, there is some moisture several feet below the surface; this moisture tends to help the *deep* insects produce more

offspring than *shallows* do. On the other hand, too much water in a patch of soil can cause some insects in it to drown. It occasionally rains, causing a great deal of water to collect in the areas favored by *deeps*.

Which type is fitter, *deep* or *shallow*? Which has greater probabilities of having more descendants? During a dry period, *deep* would seem, intuitively, to have a greater probability of reproductive success. Yet because *deeps* are more likely than *shallows* to drown during rain, one can easily imagine that over a longer interval which included both dry periods and occasional rain, the *shallow* type might have a greater probability of success. Insects can reproduce and die on a very short time scale. Is rain relevant to natural selection on this population if it will not rain for another 100,000 generations? 1,000? 10? These are questions about how wide of a range of conditions (alternatively: which subenvironments) should be considered to affect the reproductive probabilities of the population. The answers need not be obvious—even if we knew exactly when and how it would rain. We simply don't have clear criteria for deciding. Moreover, the question of whether rain counts as part of the insects' environment is surely only one of many questions about whether particular conditions which will eventually affect the insect population are among those making up the whole environment.

A characterization of fitness should tell us what it is for one type to be fittest in a given environment; by itself it doesn't tell us what counts as the environment. Yet as the preceding example shows, there are cases it is not clear how much of the world around a population should count as part of its environment. As a result, it may not be clear which types are fittest, even given a characterization of fitness

and complete information about what conditions the population would eventually experience. A theory of fitness should thus be supplemented by an answer to the following question:

For a given population, what principles determine the ranges of conditions which give rise to the reproductively relevant probabilities constitutive of biological fitness for members of a population?

The problem of the reference environment which I'll address is the absence of a clear answer to this question.

4. Environments relevant to evolution over an interval

For an initial period of time the entire population of insects can be considered to be in an arid whole environment without rain. Since this initial period is included in a longer period during which rain occurs, the population can also be viewed as being in a whole environment with occasional rain. (This is so even over the course of the initial, rain-free period, which after all is just part of the longer period.) This second environment includes a wider range of conditions than the first, in the sense that the conditions included in the first environment are a proper subset of those in the second: The first includes only aridity, while the second includes both aridity and rain. Let's make these intuitions more systematic.

In general, we can take a population of organisms at a starting time t_0 to reside in series of nested candidate (whole) environments E_n , each persisting longer than preceding ones. Each candidate environment

includes the conditions experienced by the population from t_0 until a later time t_n .⁸ For a given population and starting time t_0 , E_n 's may include wider and wider ranges of conditions as n is increased, since as time goes on the population may experience conditions it had not yet experienced. Thus E_{20} —which lasts from t_0 until t_{20} —includes the conditions included in E_{19} —those experienced from t_0 until t_{19} —along with any new conditions experienced between t_{19} and t_{20} .

This characterization suggests an obvious response to the problem of the reference environment when the effect to be explained is this: The change (or lack thereof) in frequencies of members of a set of incompatible types A_i over the course of the interval from t_0 to t_n , given that all of those types exist over that interval.

For questions about *selection over time interval T_n on types A_i present over that interval*, the relevant environment is the *interval-relative environment E_n* that encompasses conditions experienced by the population during T_n .

Note that there is nothing essentially epistemological or subjective about interval-relative environments. Each is objective, though explanatorily, causally, and predictively relevant to a different set of effects specified in terms of different time periods.

(Given my restriction in Section 2.4 on uses of “population”, the specification of the relevant interval-relative environment can affect which organisms should ideally be included in a population. Suppose that there’s a significant chance that there will be recombination between members of the population and a separate lineage L after, but not before t_m . Then organisms in L would have to be included in the

population for an interval ending after t_m , but not necessarily for an interval ending at or before t_m . In the account of extinction environments below, this will sometimes create an interaction between what should strictly be included in the population and what counts as the relevant environment.)

The notion of an interval-relative environment also provides a foundation for a solution to a more difficult reference environment problem which is needed to account for some kinds of long-term evolution; I discuss this next.

5. Environments relevant to extinction

5.1. WHY EXTINCTION ENVIRONMENTS MATTER

Lewontin (1966) argued that a population of organisms can be viewed as a sort of detector for environmental conditions as natural selection gradually modifies the population's genetic makeup. However, conditions which affected a population in the past may be "forgotten" when their effects are lost due to drift or selection in response to new conditions.⁹

Suppose you're interested in the evolution of a particular variety of protein in plants over a very long period of time. The protein may have been evolving for millennia, across many different species, and across what would intuitively be considered different environments. The interval-relative environment for such a long period of time might include an enormously wide range of conditions. It seems misleading to

say that this entire range of conditions is always relevant to selection over the entire period of the protein's evolution. Not every condition experienced by a population during a long time interval is relevant to every change in it. Some conditions cease to affect it and may be subsequently "forgotten".

Long-term evolution typically involves a series of episodes of selection. For example, at one time, a population might include alleles A and B for protein variants. A goes extinct, leaving only B present. At a later time mutation introduces a fitter allele C , and B goes extinct. Allele D arises and pushes out C , and so on. Now, in what sense must environmental conditions relevant to selection on C and D matter for selection on the earlier variants A and B ? Surely conditions occurring only long after a type has gone extinct make no difference to selection on it; likewise for conditions occurring only long before the type arises.

The conditions relevant to selection on A and B are, roughly, those conditions experienced by the population *while A and B are present in the population*; the conditions relevant to selection on C and D are roughly those experienced while *they* are present. Conditions occurring only long before or after the time in which a set of types is present in the population are causally irrelevant to their evolution. In order to characterize environments involved in many kinds of incremental phenotypic change, we need a conception of the environment relevant to selection on variation in a population while such variation persists, that is, until the time that one of the alternative types goes extinct. We need a characterization of what I'll call the "*environment relevant to extinction*", or more succinctly, the "*extinction environment*". Note,

however, that here the effect to be explained is simply extinction of one of the types, whenever that occurs; a time interval is not implicit in the explanandum.

5.2. WHICH INTERVAL-RELATIVE ENVIRONMENT?

Though the explanandum for an extinction environment does not involve a time interval, perhaps the extinction environment for a set of incompatible types is the interval-relative environment lasting until one of the types becomes extinct. This has the advantage of making the probabilities relevant to extinction depend on all conditions occurring during the interval leading up to extinction, but on none occurring only later. Thus in the desert insects example, if the *shallows* will go extinct at t_n and there will be no rain until a later time t_m , the possibility of rain is irrelevant to the probabilities concerning extinction of the *shallows*. The extinction environment would be E_n , which does not include conditions associated with rain.

Though this proposal is on the right track, it has a problem. The extinction environment *is* equivalent to one particular interval-relative environment, I'll argue. However, it will not work to specify the interval as the one lasting until the first extinction. The problem is that the time of occurrence of extinction is probabilistic, as I argue below; there is no guarantee as to when one of the alternative types will go extinct. It may be that it is most probable that one of the types will go extinct first at time t_n , but this is only the most probable outcome. The first extinction might occur at an earlier or later time. So there is no one

interval of time which is obviously relevant to extinction. Moreover, the probabilities of extinction occurring at various times may be different relative to different interval-relative environments.¹⁰

Still, an interval-relative environment E_n may be irrelevant to extinction in the sense that it persists well beyond a time at which it's overwhelmingly probable, relative to earlier environments E_m , that extinction will have occurred. On the other hand, if an interval-relative environment E_k generates a significant probability that extinction won't have occurred when it ends, conditions which occur later might play a significant role in the evolution of the population. In that case a longer-lasting interval-relative environment is more appropriately considered the one that's relevant to extinction.

I'll propose below that the extinction environment is the briefest interval-relative environment relative to which extinction is overwhelmingly probable. This environment includes conditions making extinction very probable while excluding conditions which will occur only after the time by which extinction is nearly certain. Note that the extinction environment is still an interval-relative environment on this account; however, the interval is specified in such a way as to take account of the probabilistic nature of extinction and the environment-relativity of the probabilities.

It's crucial to this approach that each interval-relative environment determine a probability (perhaps zero) that extinction of some type will occur. Some readers may think it obvious that each interval-relative environment does this; others will not. The next section is designed to convince readers of the second sort that given reasonable assump-

tions, it's plausible that each interval-relative environment determines a probability that extinction will occur. The section following the next one will spell out my proposed account of the extinction environment in more detail.

5.3. POPULATION-LEVEL TRANSITION PROBABILITIES

In this section I'll argue that a given interval-relative environment determines probabilities of extinction. In order to do that, I'll argue for the following two points about biological probabilities:

1. *1-generation individual-level* \rightarrow *1-generation population-level*:

Probabilities of reproductive success of *individual* organisms (i.e. instantiations of specified organism types) mathematically determine probabilities of states of a *population* in the next generation.

2. *1-generation population-level* \rightarrow *multi-generation population-level*:

Probabilities of states of a population in each *next* generation mathematically determine probabilities of states of a population *many* generations later.

Note that different interval-relative environments may therefore determine different multiple-generation population-level probabilities by determining different single-generation individual-level probabilities.

Some qualifications: First, those who think that there are probabilities concerning changes in the makeup of populations but not probabilities concerning numbers of offspring for individual organisms have the option of ignoring the first point. Such readers may wish

to skip to the last few paragraphs of this section. Second, I take as my starting point fitness defined by probabilities concerning organisms in the *next* generation. However, it may turn out that fitness must sometimes be defined in terms of descendants two or more generations later. (Consider a parent who has so many offspring that they exhaust the available food, leaving none for their own offspring.) My points can easily be generalized to iterations of multiple-generation reproductive cycles instead of single generations if the basic probabilities relevant to natural selection concern reproductive success over two or some other small number of generations.

Assume that a given interval-relative environment E_n determines probabilities for organisms of various types to have various numbers of offspring. (Other interval-relative environments E_m may of course determine different probabilities.) I'll argue that these individual-level probabilities determined by E_n , along with facts about frequencies in the population, determine the probabilities that the population will contain biological types in various frequencies in the next generation. Thus, for example, given a particular population made up of organisms with genotypes A and B , there will be a certain probability that in the next generation all members of the population will have the A type, a certain probability that exactly one will have B while the rest have A , a certain probability that half will have A and half B , and so on. Similar ideas can be found in population biology (e.g. Roughgarden, 1979; Ewens, 2004) and in the theory of random processes (e.g. Bharucha-Reid, 1960; Grimmett and Stirzacker, 1992). To motivate such claims, I want to show how, in one simple case, the probabilities of various states

in a population in the next generation can be derived from individual reproductive probabilities.

Thus consider a population of asexual organisms of two types A and B , all of which reproduce seasonally according to the same schedule. Assume that succeeding generations do not overlap and that the population size N remains constant. Assume further that each A always produces some particular number of offspring W_A , and each B always produces a particular number of offspring W_B ; we can take these numbers to be fitnesses. Finally, let i range over possible numbers of A 's in a given generation at t_0 , and let j range over possible numbers of A 's in the next generation at t_1 .

Of the N “slots” for possible organisms in the next generation at t_1 , each has a certain probability p of being filled by the progeny of an A parent and thus of containing an A . Each also has a probability $q = 1 - p$ of being filled by the progeny of a B parent, thus containing a B . Given reasonable assumptions, these probabilities p and q will depend on the number of A and B parents in the population, and on the probability of a given A or B parent contributing an offspring to the next generation. To see this, note that one can view parents as contributing offspring to a pool of candidates for membership in the next generation. The percentage p of A 's in the pool is then equal to the number of A parents times the number of offspring per A , divided by W_T , the total number of offspring in the pool. The percentage q of B 's in the pool is calculated similarly, with $N - i$ instead of i :

$$p = \frac{W_A i}{W_T} , \quad q = \frac{W_B (N - i)}{W_T} ,$$

where

$$W_T = W_A i + W_B (N - i)$$

is the total number of offspring in the pool. If we further make the reasonable assumption that each offspring in the pool has an equal chance of making it into any given slot in the next generation, p and q become probabilities of an A or a B , respectively, filling a slot.¹¹

We can then ask, given that the population at generation t_0 has i members of the A type, what is the transition probability P_{ij} that in the next generation, at t_1 , the population will have exactly j members of the A type? First note that P_{ij} is the probability of all choices of N individuals from the offspring pool in which an A is chosen exactly j times. Any one such set has probability

$$p^j q^{N-j},$$

since in such a set an A is chosen j times and a B is chosen $N - j$ times, and those events have probabilities p and q , respectively. (I assume samples are independent.) There are many such sets containing different individuals, so we have to multiply $p^j q^{N-j}$ by the number of these sets, which is a number represented by the expression $\binom{N}{j} = N!/(j!(N-j)!)$. The probability of producing a population of size N containing exactly j instances of A from a population of the same size with exactly i instances of A is thus a binomial distribution:

$$\begin{aligned} P_{ij} &= \binom{N}{j} p^j q^{N-j} \\ &= \binom{N}{j} \left(\frac{W_A i}{W_T} \right)^j \left(\frac{W_B (N - i)}{W_T} \right)^{N-j}. \end{aligned}$$

We have thereby derived the probability of a frequency of a type in the next generation from (a) individual-level fitnesses, and (b) frequencies of types in the current generation.¹²

Most real world biological cases are more complex, of course. In addition to the assumptions mentioned in the example, my model (a) assumed that the fitness of a given biological type in the current generation was not affected by niche construction or other effects so as to be sensitive to frequencies of other organisms in a population, their spatial locations, or other relationships between individual organisms. I also I ignored (b) cross-generational effects on fitnesses of organisms, e.g. via maternal investment in offspring or niche construction. Relaxing assumptions like (a) would mean that the reproductive probabilities of an organism depended on properties of contemporaneous members of its population. Relaxing assumptions like (b) would mean that the future proliferation of a type could depend on properties of any current or recent member of the population. Similarly, sexual reproduction would mean that the success of a type depended on its success at producing grandoffspring, which in turn depends on potential mates in the population. Non-random mating would alter these probabilities further. Other factors can affect probabilities of long-term success.

Nevertheless, it appears that in general more realistic models merely involve various failures of independence of reproductive probabilities for organisms in nearby generations. The theory of random processes includes mathematical techniques suitable for dealing with some of the relevant complexities (e.g. Bharucha-Reid, 1960), but real world cases would quickly become intractable without simplifying assump-

tions. However, such difficulties appear to be due only to limitations on computational ability and mathematical knowledge. It seems plausible that in general transition probabilities over one generation (or possibly a small number of generations) for a given population are determined by individual-level probabilities along with frequencies.¹³

Note, now, that all of the mathematical propositions presented or alluded to above are derived from probabilities determined by a particular interval-relative environment E_n . Since for a given population and starting time t_0 , each interval-relative environment can determine different probabilities of reproduction for different individual types, each environment can also determine different single-generation (or few-generation) population-level transition probabilities.

However, if there are determinate probabilities of transition from a population in a state i to various states in the next generation (or a near generation), these probabilities will mathematically imply a probability of transition from a population in state i to a population in a state k in an even later generation. For example, to calculate the probability that a population with i A 's will produce a population with k A 's two generations later, we can use the product $P_{ij} \times P'_{jk}$, which is the probability of a population in state i at t_0 going into j at t_1 and subsequently into k at t_2 . (In realistic cases, the set of transition probabilities P_{ij} from the first generation to the second need not be the same as that from the second to the third, P'_{jk} .) The probability of a population in state i going into k two generations later via *any* intermediate state at t_1 is then equal to the sum of probabilities $P_{ij} \times P'_{jk}$ over possible

intermediate frequencies j , from 0 to N :

$$\sum_{j=0}^N P_{ij} P'_{jk} ,$$

Iterating this calculation for a population in a given initial state at t_0 , relative to a given interval-relative environment E_n , there will be determinate probabilities that the population will have various states in all future generations while that environment exists.

Among these states in future generations are those in which one or the other of the alternative types A_i goes extinct. The probability of extinction for a given interval-relative environment is then the sum of the probabilities of each type going extinct in each number of generations which might occur during the interval in question.

5.4. THE EXTINCTION ENVIRONMENT

As pointed out above (§4), a population at t_0 is in many different interval-relative environments E_n which extend various lengths of time into the future. Moreover, for a given population and starting time t_0 , environments with shorter duration often encompass a narrower range of conditions than environments with longer duration beginning from the same time (and never encompass a broader range of conditions).

Some interval-relative environments E_n will determine that there is an extremely high probability of extinction before E_n ceases at time t_n . Let us call such environments “overwhelmingly-probable-extinction” (OPE) environments. I’ll stipulate that the high probability must be at least .95, but a different number could be used. Of those OPE environments, one, E , extends the least far into the future from our

starting point t_0 . I propose that this be considered the environment relevant to extinction:

The extinction environment is the OPE environment E with the shortest duration.

Why? First, an environment with a duration shorter than E 's will have a significant probability of ceasing while natural selection on the set of alternative types A_i is still possible. More precisely, relative to this shorter environment there will be a significant chance that all of the alternative types will be present when it ends. The shorter environment thus may allow a reasonable chance that conditions outside of it will affect natural selection on the A_i 's. On the other hand, environments which last longer than E can include additional factors which might matter for selection if extinction did not occur before E ceased. That occurrence, however, has a very low probability, .05 or less. (This probability is relative to E because E affects the population before the additional conditions included in a longer-running environment can make a difference.)

For example, consider the shortest environment E in which it's overwhelmingly probable that extinction of either the *deeps* or the *shallows* will occur. It's this environment whose inclusion of rain or lack of rain makes the difference (with at least .95 probability) between *deep* or *shallow* going to fixation. Shorter environments may give a significant probability that conditions not included in the environment will help determine which type goes to fixation; longer environments can include additional conditions which have an extremely low probability (relative to E) of making a difference.

One might be uncomfortable with calling E *the* extinction environment, given that it is defined in terms of a somewhat arbitrarily chosen cutoff probability of .95. This is a reasonable concern. Strictly speaking, for a given population and time there are many extinction environments. Each is the extinction environment with the shortest extent relative to a different cutoff probability. One or another may be of more interest in a given theoretical context. (This a somewhat messy answer to the question of what environment is relevant to extinction, though an advance nevertheless.) In any event, in the rest of the paper I'll continue to focus on environments relevant to extinction defined in terms of a single cutoff probability.

Objection: What about cases where there is selection for a polymorphic equilibrium state of a population, i.e. a state in which two or more alternative types are maintained in a population? What is the extinction environment then?

Reply: It's important to be clear about what possible effect it is you want explained; the extinction environment is not relevant to all effects. The extinction environment is an interval-relative environment relevant to an explanation of the disappearance of a type from the population due to natural selection. Suppose, however, that a large segment of the initial series of nested interval-relative environments make maintenance of a polymorphism very probable for a long time, yet you are interested in explaining extinction at some later date. Then there are two possibilities.

One is that the environment will eventually change so as to make extinction very probable. (For example, t_0 might be in the distant past;

the extinction might have occurred recently.) That is, relative to some interval-relative environments lasting for a very long time, extinction is probable. Then the extinction environment beginning from t_0 is the shortest such interval-relative environment which makes extinction probable enough. Given the length of this extinction environment's time interval, it might include a very wide range of conditions, but it's nevertheless the environment relevant to what you wanted explained.

The other possibility is that no environmental change occurring within any reasonable time frame will make extinction probable. That is, environments relative to all intervals of time which could possibly interest us all make it probable that the competing types will persist. In that case, there is no extinction environment, at least not for practical purposes.¹⁴

Regardless, the maintenance of a polymorphism over a time period can be explained in terms of the interval-relative environment for that interval. This is so whether the effect to be explained is maintenance of the polymorphism over an extremely long period in which the environment changes a great deal, or only over a small interval of time from within that long period. Here again different explananda require different environments as explanations/causes.

One can develop additional concepts of environment for other kinds of effects, though some such concepts are less interesting than one might think. Thus for example, the *environment relevant to initiation of a stable equilibrium* (for two competing types) would be the environment with the shortest duration which makes it overwhelmingly likely that some proportion r of types strictly between 0 and 1 of will be achieved,

such that any subsequent small departure from r makes return to r likely. An environment relevant to initiation of a stable equilibrium will not be particularly interesting in cases in which, at the moment after equilibrium is initiated, a change in conditions destroys the equilibrium. (The probabilities *were* such that the equilibrium was stable: Probabilities of prolonged departures from r were low relative to the conditions included in the environment relevant to initiation of a stable equilibrium. Those conditions ceased to obtain shortly thereafter, however.) A longer interval-relative environment will be more appropriate. There is no analogous infelicity for environments relevant to extinction, because extinction is an absorbing state: Once extinction of a type occurs, it's gone until mutation or migration reintroduces it.

6. Conclusion

In the preceding pages, I've argued that a general account of fitness is incomplete if it's not supplemented with a solution to the problem of the reference environment, a general account of what environment determines fitness. I distinguished various aspects of this problem and proposed solutions to two varieties of reference environment problems.

The grain dimension of the problem of the reference environment concerns the difference between the whole environment (all of the fitness-affecting conditions experienced by a population) and various kinds of subenvironments (subsets of the whole environment). I argued that an account of fitness relevant to natural selection on a population needs

to be supplemented by a characterization of the whole environment—whether or not it's also important to characterize various subenvironments. Although there has been some discussion of principles for characterizing certain subenvironments, I'm not aware of earlier attempts to give a general characterization of the whole environment.

The explanandum property dimension of the reference environment distinguishes between different facts about the same population which may need to be explained by different environments. I argued that when the explanandum concerns the evolution of frequencies of types over a period of time in which these types are all present, the relevant environment is the interval-relative environment encompassing conditions experienced by a population over that interval of time. However, much of evolution has to do with types which come and go through extinction and mutation and other factors; I argued that a concept of an environment relevant to extinction is needed to deal with this fact. Such extinction environments are supposed to explain any one of several types going extinct whenever that occurs. I proposed that the extinction environment is the shortest interval-relative environment such that extinction of some trait is overwhelmingly probable relative to it. Long-term evolution involving a series of gradual changes can thus be driven by a series of distinct (though perhaps similar) extinction environments.

There are several areas in which future research on concepts of environment seems warranted. First, there may well be evolutionary explananda other than those mentioned here which warrant their own notions of environment. Second, there has been little work on charac-

terizing subenvironments; it's likely that there is more to say about them.

Third, there's a need for general characterizations of the internal structure of both whole environments and subenvironments. In this paper I focused on the narrow question of characterizing the set of conditions which compose the whole environment. Yet an environment is not simply a set of conditions. Two environments could include the same ranges of conditions but determine different fitness values by making different phenomena more common. I argue elsewhere (Abrams, 2007) that a characterization of biological fitness requires an account of "organism circumstance probabilities": probabilities that organisms will experience various circumstances. This suggests that environments can differ not just in the conditions which they include, but also in determining different probability distributions over various conditions experienced. Similarly, environments viewed as collections of subenvironments could differ in the probabilities they give different subenvironments. I believe that the framework presented above can be used to help systematize such ideas.¹⁵

Fourth, one might wonder whether my approach to characterizing whole environments is relevant to debates about the utility and features of various concepts of ecological niche.¹⁶ A proper discussion of niche concepts would require a separate article, but I can make a few brief remarks. On one hand, "niche" is often used as a label for a set of conditions or a region in the world which makes the persistence of members of a given species possible or probable. What is at issue is whether a population as a whole will go extinct, or whether and why

members of a given species will be found in certain locations or among certain conditions. This use of “niche” seems to answer to different theoretical needs than those to which my proposals are directed, since it doesn’t concern changes in proportions of types over time.¹⁷ On the other hand, ecologists also ask questions about how frequencies of members of different species fluctuate in a given region. Such effects are often closely analogous to natural selection. Thus it might be fruitful to investigate applications of the present approach to “populations” consisting of members of two or more species.

Finally, a note about the place of environments in a scientific ontology: Philosophers of science often claim that science tries to “cut nature at the joints”. In some cases, the joints between objects—e.g. between a solid object and surrounding air—correspond to significant differences along many dimensions of physical variation (cf. (Campbell, 1958; Wimsatt, 1974; Wimsatt, 1981)). Biological environments, on the other hand, are complex and have ill-defined boundaries, like some functionally defined entities (Wimsatt, 1974). An extinction environment is also like certain kinds of functionally defined entities in that it is defined by probabilities of effects. The idea of defining an entity in this way can be found in early functionalist theories of mental states (Putnam, 1975) and is suggested by some accounts of biological function (Abrams, 2005; Bigelow and Pargetter, 1987; Wimsatt, 1972). But extinction environments are quite unlike functionally defined entities; among other things, they are not defined by anything vaguely like the interests of organisms. Extinction environments thus may provide a novel illustration of a significant scientific category.

Notes

¹ Smith and Varzi (1999; 2002; 2006) treat “environment” as a synonym for certain senses of “niche”; I discuss the latter term briefly in Section 6.

² Some authors have argued that niche construction and other factors imply that the organism-environment boundary is not as clear as we take it to be (Laland et al., 2001; Wilson, 2005). This issue is distinct from the ones which are my focus. We can put it to one side by asking what conditions are relevant to selection on a biological type without regard to whether these conditions are inside or outside the organism. Note that some authors view conditions well within the organism boundary as bearing a relationship to other parts of the organism much like that born to environmental conditions (e.g. Sterelny and Kitcher (1988) treat an allele as part of another allele’s “environment”).

³ I want to allow the possibility that there are cases where there really is no fact of the matter about which organisms are fitter than others.

⁴ More generally we can consider the fitness of a token of a unit of selection or of a type of a unit of selection.

⁵ Likewise, different effects, perhaps involving different properties of the same object, may have different causes, which might also involve distinct properties of common objects. There has been debate about whether fitness and natural selection are in any sense causes of evolution, whether they can be used to predict evolution, or whether they only help to explain evolution, and about the character of each of these three possible relations. I want to remain neutral about the precise relationship between fitness or natural selection and evolution. In order to avoid distracting verbosity, when I say that something “explains” evolution, the reader should understand this as meaning “explains or predicts or causes”. Similar remarks apply to cognates such as “explanatory”.

⁶ This assumption is not uncontroversial, but it has a long history and broad support (e.g. Brandon, 1990; Bouchard and Rosenberg, 2004; Richardson and Burian, 1992; Sober, 2000; Weber, 2001).

⁷ This requirement can be generalized in various ways to allow natural selection on entities other than organisms or to cover non-genetic inheritance.

⁸ We could pick an earlier or later starting point t_0 if we wanted to answer a question about the evolution of a population from a different initial time.

⁹ Levins (1968, p. 11) makes similar points. See (Wimsatt, 1980) for further discussion.

¹⁰ One might think that the conditions relevant to extinction have to do with a probability distribution over extinction times and corresponding interval-relative environments. This leads to a paradox. As I argue later, the probabilities of extinction at t are the same probabilities which constitute fitness; they're relative to an environment. It seems we'd have to choose an environment in order to derive a probability distribution over all environments—which we'd need in order to first choose an environment. There is no obvious independent source for a distribution over interval-relative environments.

¹¹ This is sampling with replacement, which makes the illustration simpler and can often be given independent motivation.

¹² I characterized fitnesses in terms of numbers of offspring in order to simplify explanation of the mathematics. However, the same equations would be appropriate if litter sizes were fixed and W_A and W_B were viabilities—probabilities of reaching reproductive age—since in my example numbers of offspring merely weighted the composition of the offspring pool. My formulation here is mathematically equivalent to one in terms of products of frequencies and relative fitnesses divided by average fitness.

¹³ What is plausibly true in principle but mathematically intractable is of course of little use to empirical investigators (though not necessarily of no use—intuitions about intractable relationships can guide approximations, for example). That doesn't mean that mathematically intractable relationships are not instantiated in the world—much scientific and mathematical investigation is predicated on the assumption that they are—and it does not mean that claims about them are not testable in principle.

¹⁴ These remarks suggest that one might want something like the following explained: Extinction before time t_n , or whatever does happen before t_n if not. In this case, the relevant environment would seem to be the shortest of two environments: the extinction environment and E_n , the interval-relative environment which lasts until t_n . Why? Well, if extinction is overwhelming probable relative to some interval-relative environment E_m that ends at $t_m < t_n$, the additional conditions experienced between t_m and t_n are not sufficiently relevant to extinction, which will very probably have already occurred by t_m . On the other hand, if no interval-relative environment ending before t_n makes extinction overwhelmingly probable, E_n is the relevant environment since it includes the conditions which make a difference to evolution between t_0 and t_n .

¹⁵ A few modifications to the picture given here might be needed. For example, perhaps not all intervals should count as defining an interval-relative environment. An interval which includes rain might need to last long enough that rain comes and goes in some recurrent or systematically probabilistic pattern in order to define an environment.

¹⁶ See (Schoener, 1989; Colwell, 1992; Griesmer, 1992; Leibold, 1995), on which the remarks in the text are based. Note that “niche construction”—which occurs when an organism’s effects on its environment affect its fitness or other organisms²—encompasses much more than the construction of ecological niches (Odling-Smee et al., 2003).

¹⁷ Smith and Varzi (1999; 2002; 2006) also sometimes use “niche” to refer to aspects of the immediate surroundings of a token organism, but this use doesn’t seem designed to address issues about changes in proportions, either.

Acknowledgements

The problem addressed above derives from a challenge posed by Ruth Millikan early in my work on my dissertation (which formulated a version of the problem in a restricted context but didn't attempt to solve it). The development of the paper benefited directly from comments by Elliott Sober, Frédéric Bouchard, Phillippe Huneman, Grant Ramsey, an anonymous reviewer, and several members of the audience at a presentation at the 2006 meeting of the Florida Philosophical Association. It benefited indirectly from discussions of related issues with Bill Wimsatt, Alex Rosenberg, Ruth Millikan, Patrick Forber, Robert Brandon, Murat Aydede, and others. I'm extremely grateful for all such help. It was ideas taken from Wimsatt's (1980) which evolved in my mind for many years before finally mutating into this paper's germ.

References

- Abrams, M.: 2005, 'Teleosemantics without Natural Selection'. *Biology and Philosophy* **20**(1), 97–116.
- Abrams, M.: 2007, 'Fitness and Propensity's Annulment?'. *Biology and Philosophy* **22**, 115–130.
- Bharucha-Reid, A. T.: 1996 [1960], *Elements of the Theory of Markov Processes and their Applications*. Dover.
- Bigelow, J. and R. Pargetter: 1987, 'Functions'. *Journal of Philosophy* **84**(4), 181–196.
- Bouchard, F. and A. Rosenberg: 2004, 'Fitness, Probability, and the Principles of Natural Selection'. *British Journal for the Philosophy of Science* **55**(4), 693–712.

- Brandon, R. N.: 1990, *Adaptation and Environment*. Princeton University.
- Byerly, H. C. and R. E. Michod: 1991, 'Fitness and Evolutionary Explanation'. *Biology & Philosophy* **6**, 1–22.
- Campbell, D. T.: 1958, 'Common Fate, Similarity, and Other Indices of the Status of Aggregates of Persons as Social Entities'. *Behavioral Science* **3**, 14–25.
- Colwell, R. K.: 1992, 'Niche: A Bifurcation in the Conceptual Lineage of the Term'. in (Keller and Lloyd, 1992), pp. 241–248.
- Ewens, W. J.: 2004, *Mathematical Population Genetics, I. Theoretical Introduction*. Springer, 2nd edition.
- Gillespie, J. H.: 1998, *Population Genetics: A Concise Guide*. The Johns Hopkins University Press.
- Griesmer, J. R.: 1992, 'Niche: Historical Perspectives'. in (Keller and Lloyd, 1992), pp. 231–240.
- Grimmett, G. R. and D. R. Stirzacker: 1992, *Probability and Random Processes*. Oxford University Press, 2nd edition.
- Keller, E. F. and E. A. Lloyd (eds.): 1992, *Keywords in Evolutionary Biology*. Harvard.
- Laland, K. N., J. Odling-Smee, and M. W. Feldman: 2001, 'Niche Construction, Ecological Inheritance, and Cycles of Coningency in Evolution'. In: S. Oyama, P. E. Griffiths, and R. D. Gray (eds.): *Cycles of Contingency: Developmental Systems and Evolution*. MIT, Chapt. 10, pp. 117–126.
- Leibold, M. A.: 1995, 'The Niche Concept Revisited: Mechanistic Models and Community Context'. *Ecology* **75**(5), 1371–1382.
- Levins, R.: 1968, *Evolution in Changing Environments*. Princeton.
- Lewontin, R. C.: 1966, 'Is Nature Probable or Capricious?'. *BioScience* **16**(1), 25–27.
- Mills, S. and J. Beatty: 1979, 'The Propensity Interpretation of Fitness'. *Philosophy of Science* **46**(2), 263–286. Reprinted in (Sober, 1994).
- Odling-Smee, F. J., K. N. Laland, and M. W. Feldman: 2003, *Niche Construction: The Neglected Process in Evolution*. Princeton.
- Putnam, H.: 1975, 'The Nature of Mental States'. In: *Mind, Language and Reality*. Cambridge University.

- Ramsey, G.: 2006, 'Block Fitness'. *Studies in History and Philosophy of Biological and Biomedical Sciences* **37**(3), 484–498.
- Richardson, R. C. and R. M. Burian: 1992, 'A Defense of Propensity Interpretations of Fitness'. In: *PSA 1992*, Vol. 1. pp. 349–362.
- Rosenberg, A.: 1994, *Instrumental Biology or the Disunity of Science*. University of Chicago Press.
- Roughgarden, J.: 1979, *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Macmillan.
- Schoener, T. W.: 1989, 'The Ecological Niche'. In: J. M. Cherrett (ed.): *Ecological Concepts: The Contribution of Ecology to an Understanding of the Natural World*. Blackwell, pp. 79–113.
- Smith, B. and A. C. Varzi: 1999, 'The Niche'. *Noûs* **33**:2, 198–222.
- Smith, B. and A. C. Varzi: 2002, 'Surrounding Space: The Ontology of Organism-Environment Relations'. *Theory in Biosciences* **121**(2), 139–162.
- Smith, B. and A. C. Varzi: 2006, 'Environmental Metaphysics'. In: U. Meixner (ed.): *Metaphysics in the Post-Metaphysical Age. Proceedings of the 22nd International Wittgenstein-Symposium*. pp. 231–239.
- Sober, E.: 1984, *The Nature of Selection*. MIT Press.
- Sober, E. (ed.): 1994, *Conceptual Issues in Evolutionary Biology, 2nd ed.* MIT Press.
- Sober, E.: 2000, *Philosophy of Biology*. Westview Press, 2nd edition.
- Sterelny, K. and P. Kitcher: 1988, 'The Return of the Gene'. *Journal of Philosophy* **85**, 339–361.
- Weber, M.: 2001, 'Determinism, Realism, and Probability in Evolutionary Theory: The Pitfalls, and How to Avoid Them'. *Philosophy of Science* **68**(3), S213–S224.
- Wilson, R. A.: 2005, *Genes and the Agents of Life*. Cambridge.
- Wimsatt, W. C.: 1972, 'Teleology and the Logical Structure of Function Statements'. *Studies in History and Philosophy of Science* **3**(1), 1–80.
- Wimsatt, W. C.: 1974, 'Complexity and Organization'. In: K. F. Schaffner and R. S. Cohen (eds.): *PSA 1972*. pp. 67–86. Reprinted in (Wimsatt, 2007).
- Wimsatt, W. C.: 1980, 'Randomness and Perceived Randomness in Evolutionary Biology'. *Synthese* **43**, 287–329.

- Wimsatt, W. C.: 1981, 'Robustness, Reliability, and Overdetermination'. In: M. B. Brewer and B. E. Collins (eds.): *Scientific Inquiry and the Social Sciences*. Jossey-Bass, pp. 124–163. Reprinted in (Wimsatt, 2007).
- Wimsatt, W. C.: 2007, *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*. Harvard.

